Bioprospecting for Microorganisms and Enzymes with Biorefining Potential

by

Laura Frances Lyons

Supervisors Dr Kerrie Farrar and Dr Justin Pachebat

A thesis submitted at the Institute of Biological, Environmental and Rural Sciences (Aberystwyth University), for the degree of Doctor of Philosophy.

2015

Declaration

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed ..... (candidate)

Date .....

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Where *correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

Other sources are acknowledged by footnotes giving explicit references.

A bibliography is appended.

Signed ..... (candidate)

Date .....

[*this refers to the extent to which the text has been corrected by others]

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... (candidate)

Date .....

Word count of thesis: 53 812

Acknowledgments

Well here's the cheesey part (mmm cheese, sorry, let us continue…). Firstly, I would like to thank my supervisors Kerrie and Justin, especially for giving me a kick up the arse when I needed it. Also, it can't be easy reading draft after draft of long winded explanations of relatively simple things. Massive thanks must go out to Sarah Hawkins, who has always been there in times of trouble (like the hot water baths…) as well as when things were actually going smoothly (I'm sure that must have happened once…). Personally, I think every lab should have a Sarah! Also, thank you to Naomi (along with Sarah) for her patience in training me in the first year. I would like to thank everyone in the labs I have worked in, except Abhi as he stole everything; didn't make up buffers when he'd used them up, and generally left a mess everywhere (just kidding). Everyone in the student offices over the past four years obviously deserves thanks. Especially Magda, who brought me back down to earth when everything was going wrong, and Alice for her distractions of field work, as well as maintaining a working atmosphere in the office. Not forgetting Dave and Tim for the endless football discusions and Fantasy Football rivalry. Naheed and Colin must also be acknowledged for their great knowledge of proteins and patience when teaching me. Thank you to David Bryant and Naheed for also reading through chapters and making me think outside of the box. Also, thanks to Joe, Ana, Richard and Eli for not talking about work on Friday evenings down The Glen. And Jenny, Tim, Dave and Alina for all the fun times! Plus, Mark, Jan, Pip and Liz for the rugby times! And of course Spicer with her lunchtime walks with MollyDog! I must also acknowledge Manny (Romain Jugal), for his hard work and dedication in the six months he worked with me as part of his MSc. The lab was not the same after Christmas songs in March! (Où est la piscine?!)

I'd like to also extend thanks to people outside of academia who have encouraged and congratulated me over the extent of my PhD, including Andy, Nige, Olive, Iris, Kellie, Marc and many more. The Sunday quiz in Scholars was always a good end to the week, thank you Tom. Of course, I thank my family for their endless support, especially my mom and dad for the gift of rent for Christmas and my birthday! And my brother, who occasionally knows when he is wrong, but has always encouraged his little sister. Mostly, I would like to thank Grace for putting up with me, supporting me and generally just trying to make life a bit simpler and easier.

Finally I would like to thank my funders because without them I would not have had this amazing experience and potentially helped the human race (yes, I can still change the world even after doing a PhD!). Also, thank you for reading this! I hope it's not too tedious and I apologise, I'm sure most of the pages are pictures!

Summary

The rise of biorefining and application of biotechnology to combat climate change and accomplish energy security is a necessity. There have already been steps to produce sustainable biofuels and products throughout the world. However, not all processes are economically viable due to costs of enzymes, pre-treatments, and scale-ups. In this project *Miscanthus sp.* was the main source of bacterial isolation due to its bioenergy characteristics as a low-input high-output crop. Specifically, due to the high sugar content of harvested senesced *Miscanthus*. The aim of this project was to discover novel microbes and enzymes with potential to contribute to the generation of fuels and chemicals from plant biomass. Specifically, we aimed to isolate and characterise enzymes capable of efficiently releasing sugars from pre-treated lignocellulosic biomass, for subsequent fermentation to ethanol and platform chemicals.

Aerobic bacteria were cultured from harvested chipped *Miscanthus* and soil surrounding *Miscanthus* crops and were characterised morphologically, functionally and taxonomically. Bacteria in our collection included amongst others: *Pseudomonas sp.*, *Burkholderia sp.*, *Variovorax paradoxus*, *Luteibacter sp.* and *Bacillus sp..* The collection was screened for carbon utilisation using cellulose (in the form of carboxymethyl cellulose), xylan (from beechwood) and starch by enzymatic activity, at a range of temperatures and pHs. From the bacterial library, 88.5% of cultures showed cellulase activity, 93.2% xylanase activity, 79.7% starch degradation activity over the temperature or pH range, with 66.2% demonstrating activity over all three assays. Proteins were isolated from bacteria that demonstrate effective starch utilisation for further characterisation. Bacterial isolates that exhibited xylan utilisation at high pH and temperature were characterised by whole genome sequencing to identify interesting enzymes and pathways using bioinformatics software CLC genomics and Seed RAST. Finally, homologous proteins have been modelled using Phyre2 and 3DLigandSite to analyse structure and binding sites.

This work was part of the wider BEACON project which aimed to establish Wales as a Biorefining Centre of Excellence. BEACON built integrated ‘Green Supply Chains’ with a focus on developing new routes to functional, cost competitive products using biomass rather than fossil fuels.

Contents

List of Abbreviations

A	Adenine
Abf2	Arabinosidase 2
ACS	American Chemical Society
AFEX	Ammonium fibre explosion
AmyA	Alpha-amylase gene
APS	Ammonium persulfate
A.T.	Arabinose transporter
ATCC	American Type Culture Collection
ATM	Amplicon Tagment Mix
BAC	Bacterial Artificial Chromosome
BBH	Bi-directional best hit
BBSRC	Biotechnology and Biological Sciences Research Council
BCIP	Bromochloroindolyl phosphate
β-Glr	β-glucuronidase
BG	β-glucosidase
BHU(2)	Biomass Hydrolysis Unit
BLAST	Basic Local Alignment Search Tool
βME	Beta-mercaptoethanol
bp	Basepair
BRENDA	BRaunschweig ENzyme DAtabase
BSA	Bovine Serum Albumin
BX	β-xylosidase
C	Cytosine
CAZy	Carbohydrate Active enZYmes
CBCB	Centre for Bioinformatics and Computational Biology
CBH1 and CBH2	Cellobiohydrolase 1 and 2
CBM	Cellulose Basal Medium
CBP	Consolidated Bioprocessing
CDD	Conserved Domain Database
Cel5A	Endo-β-1,4-glucanase 2
Cel61A	Endo-β-1,4-glucanase 4
Cel12A	Endo-β-1,4,glucanase 3
CMC	Carboxymethyl cellulose
C.T.	Cellobiose/cellodextrin transporter
CTAB	Hexadecyltrimethylammonium bromide
CUAs	Carbon Utilisation Assays
C3	Carbon fixation in plants, C3 uses a 3-carbon molecule in the first product of fixation
C4	Carbon fixation in plants, C4 uses a 4-carbon molecule in the first product of fixation
C5	Pentose sugars
C6	Hexose sugars
DES	DNase/Pyrogen-Free Water
dH_2O	Distilled/dionised water
DMF	Dimethylformamide
DNA	Deoxyribonucleic Acid

dNTPs	Deoxynucleotide Solution Mix
dsDNA	Double stranded Deoxyribonucleic Acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EG1	Endo-β-1,4-glucanase
EPOBIO	Realising the Economic Potential of Sustainable Resources – Bioproducts from Non-Food Crops
ExPASy	Expert Protein Analysis System
EX2	Endo-β-1,4,-xylanase 2
EX3	Endo-β-1,4-xylanase 3
FIS	Factor for inversion stimulation
G	Guanine
gDNA	Genomic DNA
GENPLAT	GLBRC (Great Lakes Bioenergy Research Centre) Enzyme Platform
GH	Glycosyl Hydrolase
GH61	Glycosyl Hydrolase family 61
GMO	Genetically Modified Organism
g/L/h	Grams per litre per hour
G.T.	Glucose transporter
His	Histidine
H-NS	Histone-like nucleotide structuring protein
HPLC	High performance liquid chromatography
HS	High sensitivity
ID	Identification
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kb	Kilobases
KAAS	KEGG Automatic Annotation Server
KEGG	Kyoto Encyclopedia of Genes and Genomes
LB	Luria-Bertani
LBM	Lignin modifying enzymes Basal Medium
Mbp	Million base pairs
MEGA	Molecular Evolutionary Genetics Analysis
MRS	De Man, Rogosa and Sharpe
MW	Molecular weight
NBT	Nitro Blue Tetrazolium
NCBI	National Centre for Biotechnology Information
ND	No data found/Not determined
Ni-NTA	Nickel-Nitrilotriacetic acid
NPM	Nextera PCR Master Mix
NR	Not reported
NT	Neutralize Tagment Buffer
OD	Optical density
ORF	Open Reading Frame
o/n	Overnight
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction

PDB	Protein Data Bank
PEG	Polyethylene glycol
pI	Isoelectric point
POP	Performance Optimated Polymer
PPP	Pentose phosphate pathway
PPS	Protein Precipitation Solution
PSI-BLAST	Position-Specific Iterated Basic Local Alignment Search Tool
RAST	Rapid Annotation of microbial genomes using Subsystems Technology
RCSB	Research Collaboratory for Structural Bioinformatics
RDP	Ribosomal Database Project
RFS	Renewable Fuels Standard
rpm	Rotations per minute
RNA	Ribonucleic Acid
rRNA	Ribosomal Ribonucleic Acid
RSB	Resuspension Buffer
SC-CO_2	Supercritical CO_2
SDS	Sodium dodecyl sulphate
SEWS-M	Salt/ethanol wash solution, DNase-free
SHF	Separate Hydrolysis and Fermentation
SIB	Swiss Institute of Bioinformatics
SNP	Single Nucleotide Polymorphism
SOB	Super-optimal broth
SOC	Super-optimal broth catabolite repression
SSF	Simultaneous Saccharification and Fermentation
SSCF	Simultaneous Saccharification and Co-fermentation
T	Thymine
TAE	Tris-acetate-EDTA
TBS	Tris-Buffered Saline
TBST	Tris-Buffered Saline plus Tween 20
TCA	Trichloroacetic acid
TD	Tagment DNA Buffer
TE	Tris-EDTA
Temed	N,N,N',N'-Tetramethyl-ethylenediamine
U.S.	United States (of America)
UV	Ultraviolet
XBM	Xylan Basal Medium
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside
X.T	Xylose transporter
xylA	Xylose isomerase
XYL1	D-xylose reductase
XYL2	Xylitol dehydrogenase
YM	Yeast and mould
3D	Three dimensional
3-HPA	3-Hydroxypropionic acid
16S rRNA	A constituent of the 30S small subunit of prokaryotic ribosomes

List of Figures

List of Tables

Chapter 1 Introduction

1.1 The problem: Climate change

Human induced climate change through the use of fossil fuels, which release greenhouse gasses into the environment, is a very real problem. Raised sea levels, drier soils, greater weather extremes and warmer temperatures have all been predicted (Adger, Huq, Brown, Conway, & Hulme, 2003; Bulkeley & Newell, 2015; Cotula, Dyer, & Vermeulen, 2008; Danielsen *et al.*, 2009). Changing from fossil fuels to a more sustainable resource, while still having access to the remaining resources, should decrease manmade emissions as knowledge and efficiency increases (Department of Energy & Climate Change, 2012). Optimisation of processing new feedstocks is required to gain the highest yield of products by utilising the available biomass, such as senesced *Miscanthus*. Bioprospecting is the discovery, collection and deriving of genetic material from samples of biodiversity that can then be used in a range of industrial applications (Pan, 2006). Here it is applied to the activities of microorganisms in degrading plant matter to produce sustainable products such as: pharmaceuticals, chemicals, fuels, cosmetics, textiles, heat, and power (BEACON, 2011).

1.2 A solution: Biorefining

Biorefining is the sustainable integration of biomass conversion processes and equipment to produce fuels, heat and value added chemicals from biomass, which is analogous to petroleum refinery (Langeveld, de Jong, van Ree, & Sanders, 2010). Being carbon-neutral and renewable, biomass is appealing as a substitute to current fossil fuels, especially petrol which can be replaced by bioethanol and biobutanol (Nakayama, Kiyoshi, Kadokura, & Nakazato, 2011; Yoshida *et al.*, 2008). This project

fitted into the secondary processing aspect of the BEACON pipeline (Figure 1.1) and depolymerisation to form an intermediate platform (Figure 1.2).

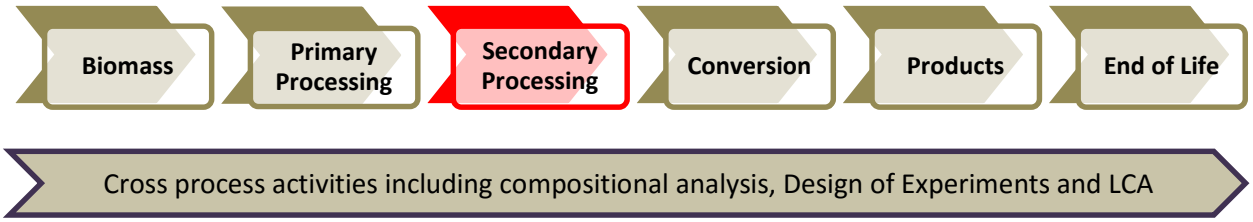


Figure 1.1 Schematic of the BEACON pipeline highlighting secondary processing which this project encompassed.

Secondary processing involves plant cell wall degradation to monomeric sugars. Thus leading into conversion, which comprises microbial metabolism of the biomass into useful chemicals, this can then be converted into products. *LCA – Life Cycle Analysis.*

Biorefining has emerged to replace petroleum refinery for sustainable energy production. Many industrial biobased products result from physical or chemical treatment and processing of biomass e.g. starch, lignin, oil, protein, cellulose and terpene (Kamm & Kamm, 2004b). Using biotechnological methods other chemicals can be produced from feedstocks, such as ethanol, lactic acid, acetone, butanol, and amino acids (H. Cheng & Wang, 2013). Feedstocks for biorefinery can include dedicated energy crops (such as *Miscanthus*), trees, aquatic plants, wood residues, waste materials and agricultural food and feed crop residues (Chin & H'ng, 2013; Kamm & Kamm, 2004b). Waste materials can include waste biomass from landscape cultivation as valuable organic compounds can be extracted from this raw material (Kamm & Kamm, 2004b). Biorefineries convert biological raw materials into final products and industrial intermediates via technological combinations, including physical, chemical and biological mechanisms (Kamm & Kamm, 2004a). However, most feedstocks require pre-treatment as saccharification can only occur when the

lignin/hemicellulose/cellulose matrix has been loosened (Chin & H'ng, 2013; Kamm & Kamm, 2004b). There are many ways to do this which will be covered in section 1.3.5 but most involve thermal, thermo-mechanical or thermo-chemical treatment which requires energy, which in turn raises costs (Kamm & Kamm, 2004b). Ultimately, biorefinery aims to achieve efficient fractionation of lignocellulosic biomass, which can then be directed into multiple streams to isolate value-added compounds so that the economics of biofuel production can be improved significantly (B. Wang, Wang, & Feng, 2010).

There are four types of bioconversion process configurations; these include Separate Hydrolysis and Fermentation (SHF), Simultaneous Saccharification and Fermentation (SSF), Simultaneous Saccharification and Co-fermentation (SSCF) and Consolidated Bioprocessing (CBP). SHF prevents problems in fermentation or hydrolysis affecting other steps and allows the microbes and enzymes to operate at optimum conditions (H. Cheng & Wang, 2013). However SSF allows both the enzymatic hydrolysis and fermentation to occur at the same time, reducing the time-scale (Ingram *et al.*, 1998). This can also eliminate problems of inhibition of cellulases by hydrolysis products (glucose and short chain cellulose chains) by fermenting glucose into ethanol when it appears in solution (H. Cheng & Wang, 2013).

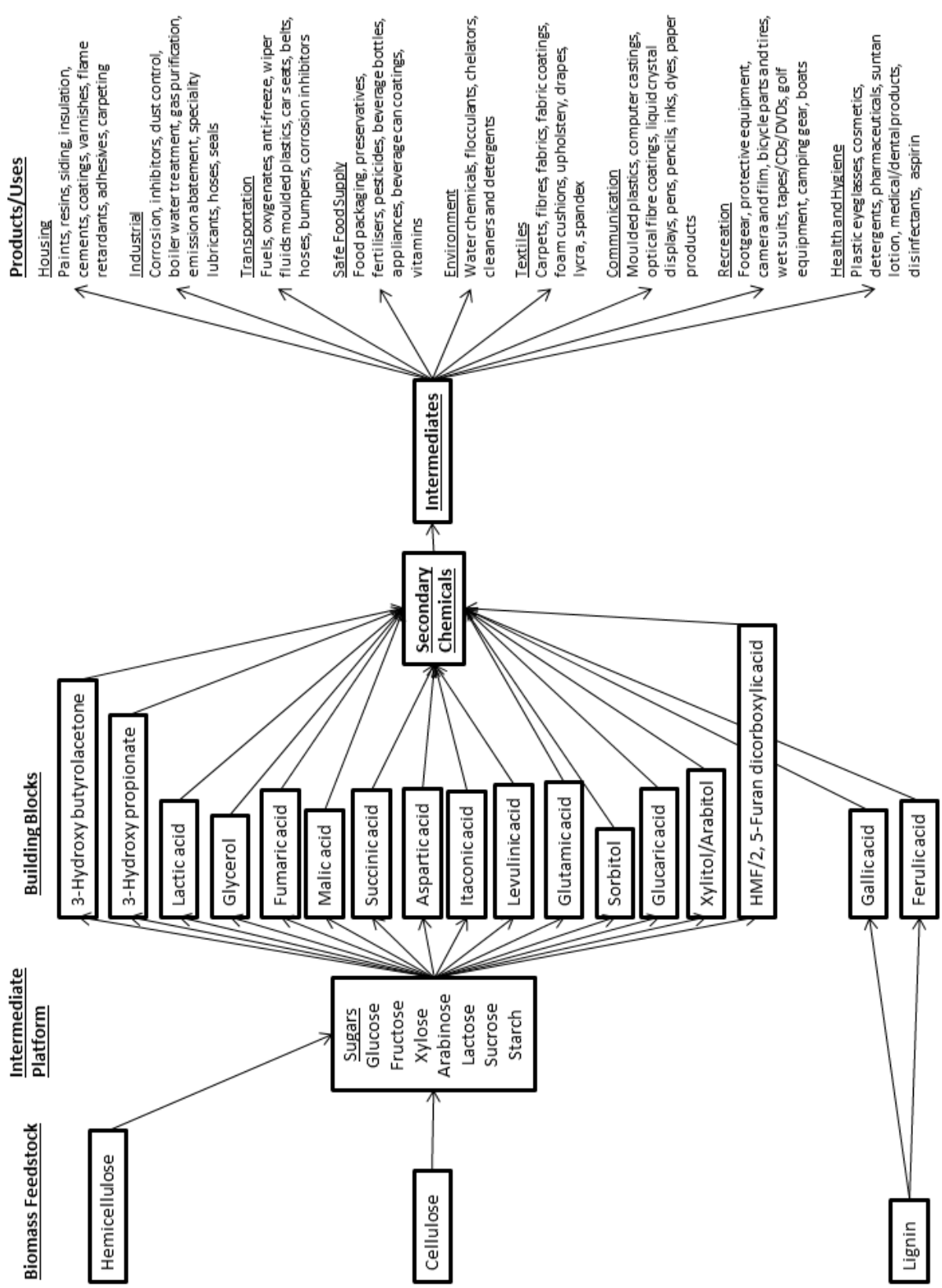


Figure 1.2 Adapted flow diagram of value added chemicals from sugars and lignin, and their final products and uses (Werpy & Petersen, 2004).

Conversion of lignocellulose to bioethanol using SSF or SSCF may help to prevent product inhibition by insufficient β-glucosidase activity. This causes a build-up of cellobiose, which means endoglucanase and exoglucanase activities are inhibited (Jeng *et al.*, 2011). However, high ethanol concentrations can also inhibit the microbe and enzyme, though this is less problematic than high concentrations of hydrolysis products. Operating temperatures for SSF (37-40 °C) can also cause problems as different cellulases have different optimal temperatures (usually between 45-50 °C), as do the microorganism used for fermentation (usually 28-35 °C) (H. Cheng & Wang, 2013). However, a simple solution to such a problem is to use a microbe with a higher optimal temperature that matches the optima of the enzymes used or vice versa. Microbes and enzymes isolated in this project could therefore aid these types of bioconversion (section 3.2). SSCF allows a microbe to ferment both hexose and pentose sugars for the production of e.g. bioethanol (H. Cheng & Wang, 2013) in a single step, especially with the use of soaking in aqueous ammonia pre-treatment whereby xylose does not have to be recovered separately from the liquid fraction (T. H. Kim & Lee, 2005). A number of bacteria isolated in Chapter 3 were shown to utilise both cellulose and xylan (section 3.2.5) and could therefore be used in SSCF. CBP uses an engineered strain to produce enzymes (e.g. cellulase) to hydrolyse sugars found in the biomass (e.g. cellulose) and ferment the sugars all in one step (Cairns, Gallagher, Hatch, & Humphreys, 2007; S. J. Lee, Warnick, Pattathil, Alvelo-maurosa, *et al.*, 2012; Lynd, Weimer, Zyl, & Pretorius, 2002; Tomes, Lakshmanan, & Songstad, 2011, pp. 13–15).

There are a variety of disadvantages for biorefining, the most important being reaction rate, as many biological conversions are slow and require specific conditions. Low product concentrations also lead to high product recovery costs. To become a viable alternative to petroleum refinery the biorefining process has to become cost-effective and competitive (H. Cheng & Wang, 2013). By reducing cost associated with pH and temperature control during saccharification the process could become more economically viable. For these reasons extremes of pH and high temperatures were investigated in this project (section 3.2.5.2).

1.2.1 *Miscanthus* as a biorefining feedstock

Degrading *Miscanthus* and soil surrounding *Miscanthus* plots were used as a source for bacterial isolation throughout this project due to its bioenergy crop status. The biofuel market demands great quantities of energy crops, such as *Miscanthus*, to cope with rising demands for ethanol, as agricultural crop residues such as wheat straw and corn stover are a limited resource (Birur, Hertel, & Tyner, 2008). *Miscanthus* is a C4 photosynthetic perennial, which is propagated via rhizome; these grasses have a higher lignin content and higher cellulose crystallinity (Vintila *et al.*, 2010). Sugarcane is another C4 species that is also used for bioethanol production as C4 plants are more efficient than C3 plants for nutrient, water and light conversion into harvestable biomass. Biofuel from sugarcane in Brazil proffers above 40% of the petrol demand already (Shrestha, Szaro, Bruns, & Taylor, 2011). However, sugarcane requires good quality land and fertilizer to have high yields, whereas *Miscanthus* can grow on marginal lands requiring little to no nutrient input and command high yields. For example, *M. sinensis* has an yearly yield of 20-26 tonnes (of dry weight) per hectare and *M. giganteus* yields over 30 tonnes per hectare per year in the spring harvest, and

both can grow proficiently in acid sulphate soils (Allison, Morris, Clifton-Brown, Lister, & Donnison, 2011; Yoshida *et al.*, 2008). However, unlike sugar cane *Miscanthus* contains very little water soluble carbohydrate at harvest and therefore is predominantly a lignocellulosic biofuel crop. Thus, the choice of *Miscanthus* as the feedstock for this project ensures avoidance of problems such as jeopardising food security, and carbon inefficient non-sustainable crops. It also aids the move towards lignocellulosic next generation perennial biomass crops, which are able to grow on land that is marginal for food and feed production. Nutrient recycling and carbon sequestration are also important in such crops, unlike in palm oil production, which causes a loss to major carbon sinks and biodiversity.

1.3 Hydrolysis of the plant cell wall

Three carbon sources (cellulose, xylan and starch) were utilised in this project (section 3.2.5) as they were major components in plant biomass. In bioenergy crops many sugars are found in the lignocellulosic plant cell wall. The objective of biorefining is to release these sugars to convert them into biofuels and platform chemicals. Biorefining of the plant cell wall requires saccharification, which involves the hydrolysis of structural cell wall carbohydrate polymers to their corresponding monomeric sugars. Plants have naturally evolved a recalcitrant cell wall to prevent physical, enzymatic and chemical degradation, which must be overcome for effective saccharification to be achieved (EPOBIO, 2006; J.-W. Lee *et al.*, 2008).

After growth culminates, a secondary wall is generated which contains 20-30% lignin, 20-30% hemicellulose and 40-50% cellulose, and makes up more than 95% of the cell wall material weight of wood (EPOBIO, 2006); the cell wall structure can be seen in Figure 1.3. Specifically, in the bioenergy crop *Miscanthus* species, the biomass equates

to 70-90% cell-wall, which is comprised of 43% cellulose, 24% hemicellulose and 19% lignin (Allison *et al.*, 2011; D. Lee, Owens, Boe, & Jeranyama, 2007); a comparison of different biomass feedstock compositions is shown in Table 1.1. As the secondary cell wall is synthesised, lignin saturates the primary cell wall, providing rigidity, structure, hydrophobicity, and the ability to withstand microbial infection. For this reason, the target microorganisms for discovery when bioprospecting were likely to be robust and saprophytic/phytopathogenic. Bacteria were therefore expected to have enzymes that could work with current commercial cocktails, resulting in lower enzyme loading and quicker biomass hydrolysis (King, Donnelly, Bergstrom, Walker, & Gibson, 2009).

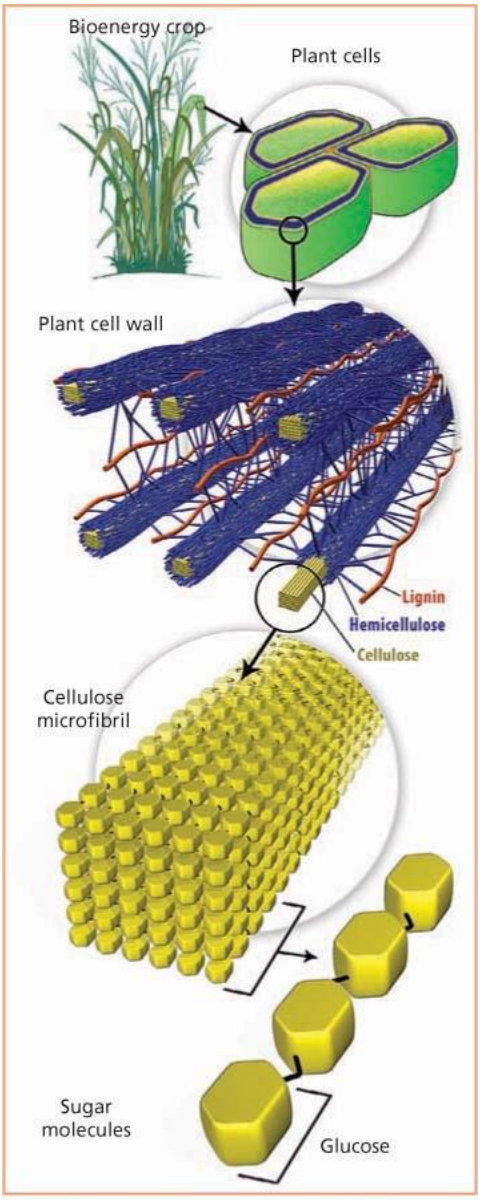


Figure 1.3 The composition of lignocellulose, from plant to single glucose monomers; taken from (Wyman & Yang, 2009).

Lignin, cellulose and hemicellulose make up lignocelluloses. Within the plant cell wall hemicellulose has a highly branched structure which binds to cellulose microfibrils. Ether bonds cross link lignin to cellulose and hemicellulose, whereas ester bonds are believed to only cross link lignin to hemicellulose, while hydrogen bonds provide intermolecular bridging between lignin and both cellulose and hemicellulose (Faulon, Carlson, & Hatcher, 1994; Harmsen, Huijgen, Bermudez Lopez, & Bakker, 2010).

Table 1.1 Monocot biomass feedstock compositions based on dry weight percentages.

Modified from EPOBIO 2006; Harmsen *et al.* 2010. *ND – no data found; * - also composed up to 35% soluble sugars (mostly fructans).*

Feedstock	Cellulose (%)	Hemicellulose (%)	Lignin (%)	Ash (%)	Protein (%)	Reference
Perennial						
Switchgrass (late cut)	44.9	31.4	12	4.6	4.5	(EPOBIO, 2006)
Miscanthus	38.2-52.1	24.8-33.9	9.2-24	2	3	(Hodgson *et al.*, 2011; D. Lee *et al.*, 2007; T De Vrije, Haas, Tan, Keijsers, & Claassen, 2002)
High sugar perennial ryegrass*	35	25	2-6	ND	ND	(Charlton, Elias, Fish, Fowler, & Gallagher, 2009)
Annual						
Corn stover	36.4	22.6	16.6	9.7	ND	(EPOBIO, 2006)
Wheat straw	38.2-44.9	24.7-27.9	8.3-23.4	6.3-10.3	1.9	(EPOBIO, 2006; Tabil, Adapa, & Kashaninejad, 2011)
Barley	33.3-46.9	20.4-27.4	11.9-17.1	2.2-6	1.1-3.6	(Adapa, Engineering, Karunakaran, Tabil, & Schoenau, 2009; Tabil *et al.*, 2011)
Sorghum	23-34	14-17	11-16	ND	ND	(D. Lee *et al.*, 2007)
Rice	33.4	28.2	7.4	ND	12.8	(Y. He, Pang, Liu, Li, & Wang, 2008)

1.3.1 Lignocellulolytic enzymes; application in a biorefining context

Lignocellulolytic enzymes are required as part of an extensive armoury to break down biomass for a number of reasons. These include exploitation of natural resources; lower carbon dioxide emissions and to create a sustainable future. Ultimately used to deliver sustainable products in a bio-based economy (Rosegrant, Zhu, Msangi, & Sulser, 2008). Although formally considered not cost-competitive, enzymatic saccharification is a promising biorefining technology. Saccharification is simply the digestion of plant cell walls into C5 and C6 sugars and is dependent on enzyme/substrate contact for adsorption. For efficient release of sugars from the plant cell wall, from biomass, a large array of enzymes are needed due to the complexity of the feedstock matrix, which causes a bottleneck for the progression of economic biorefinery processes. However, enzymes are expensive so the process needs to be cost effective and the conversion of lignocellulosic biomass should be efficient (EPOBIO, 2006). Saccharification takes place before fermentation as microbes used are generally unable to utilise anything other than monomers, such as *S. cerevisiae* (Banerjee, Scott-Craig, & Walton, 2010). Therefore enzyme cocktails are added to pre-treated biomass for depolymerisation. Cell wall saccharification efficiency is dependent upon cell wall characteristics (cellulose crystallinity, surface area, degree of polymerisation, lignin content, cellulose reactivity, degree of O-acetylation), hydrolysis conditions (enzyme mixture and enzyme type), and which pre-treatment was used. Optimisation of these factors will allow the most efficient saccharification to occur (EPOBIO, 2006).

1.3.2 Cellulose and cellulases

The main structural component of the plant cell wall is cellulose (Rubin, 2008) and is considered the most abundant biorenewable material (Swatloski, Spear, Holbrey, & Rogers, 2002). In simplistic terms, cellulose is a polymer of glucose monomers, dimers of which are termed cellobiose. The degree of polymerisation (the number of glycosyl residues per cellulose chain) affects the properties of cellulose. There can be 800 to 1700 glucose units making up one polymer molecule, which are long, straight and held together by β-1, 4 glycosidic bonds within the chain (Harmsen *et al.*, 2010). There are two types of cellulases; endocellulases, which break down cellulose randomly within the chain releasing sugars and oligosaccharides, and exocellulases, which yield cellobiose from the chain ends (Ingram *et al.*, 1998; Jeng *et al.*, 2011). Cellobiose, which is inhibitory to exocellulases, is broken down to glucose by β-glucosidase (King *et al.*, 2009). Cellulases hydrolyse cellulose at a faster rate when less crystallinity is present so softer materials are more easily converted to freely available fermentable sugars (Vintila *et al.*, 2010; Yoshida *et al.*, 2008).

Cellulose can be found in two states, crystalline and amorphous. An alignment of polymer chains creates microfibrils (Figure 1.3), which combine to form fibres (Harmsen *et al.*, 2010). Cellulose crystallinity has been found to both affect and have no effect on the saccharification rate (EPOBIO, 2006; Mansfield, Mooney, & Saddler, 1999). Crystalline cellulose is caused by inter and intra hydrogen bond formations to the microfibril via free hydroxyl groups found in the molecule, which makes it mechanically stronger and more resistant to enzymatic activity (Barnette *et al.*, 2011; Blake *et al.*, 2006; Carpita, 1996; Cosgrove, 2005; Matthews *et al.*, 2011). Amorphous cellulose is disordered and non-crystalline and lacks the mechanical structure of

crystalline cellulose (Blake *et al.*, 2006). It has been shown that amorphous cellulose can crystallise after enzymatic treatment (used to break down the cellulose) and thus a higher percentage of crystalline cellulose would be gained in the hydrolysed biomass (Behera, Arora, Nandhagopal, & Kumar, 2014; Bertran & Dale, 1985). Figure 1.4 shows cellulases attacking both amorphous and crystalline cellulose with synergistic interaction, and can therefore utilise most of the substrate. For example the exoglucanases (*CBH1* and *CBH2*) cleave the cellulose at both reducing and non-reducing ends whereas the endo and exo acting cellulases break down cellulose into fibres (Song, 2011).

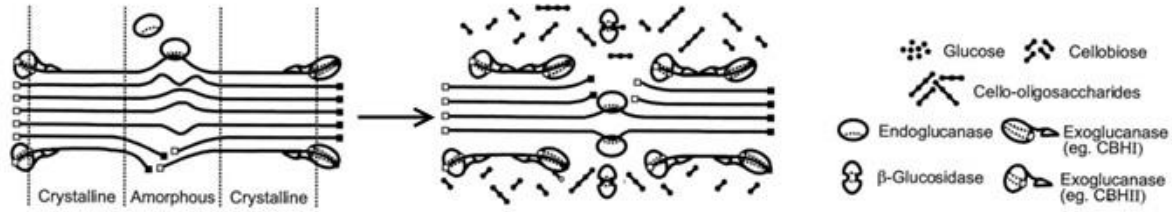


Figure 1.4 Hydrolysis of cellulose by cellulases (Song, 2011).

Cellulose hydrolysis uses a range of cellulase enzymes added as a cocktail isolated as a crude product from lignocellulosic fungi such as *Trichoderma reesi*, *Trichoderma viride* and *Aspergillus niger* (H. Cheng & Wang, 2013). Some pre-treatments degrade the biomass so that much of the cellulose can be converted. However, for commercial biorefinery pentoses need to be utilised and therefore a larger range of efficient enzymes including hemicellulases are required (Brandt, Gräsvik, Hallett, & Welton, 2013). Industrially pertinent enzymes and their substrates can be seen in

Table 1.2 (EPOBIO, 2006). Cellulases and hemicellulases are industrially sourced from the fungus *Trichoderma reesei*, although it has been shown that a more diverse enzyme mixture is required for complete biomass hydrolysis (King *et al.*, 2009). Therefore, alternative microbes need to be identified and exploited.

Biomass structure has a great influence on the ability of cellulolytic enzyme solutions to penetrate the biomatrix at different structural levels including tissue, cell and cell wall (Bryant *et al.*, 2011). The rate of hydrolysis is affected by a number of characteristics, namely particle size and specific surface, which are important for enzyme contact with the substrate and thus saccharification. Enzyme access and hydrolysis can be amplified by increased biomass fractionation, creating a higher surface area to weight ratio, allowing hydrolytic enzymes to adsorb to the substrate (da Costa Lopes *et al.*, 2013). Cellulose reactivity is also a factor affecting hydrolysis, as hydrolytic enzymes adsorb to the substrate surface, blocking hydrolytic activity and slowing the saccharification rate down (EPOBIO, 2006; B. Yang, Willies, & Wyman, 2006; B. Yang & Wyman, 2006). Studies have demonstrated that the degree of polymerisation does not affect hydrolysis (EPOBIO, 2006; Mansfield *et al.*, 1999; Sinitsyn, Gusakov, & Vlasenko, 1991). Hydrolytic enzymes have also been found to adsorb to lignin in lignocellulosic substrates; saccharification efficiency decreases as lignin content increases, although complete lignin removal is not necessary to improve degradability (EPOBIO, 2006). Therefore, energy inputs rise, due to the need to get higher temperatures or more concentrated acids/alkalis to de-lignify the material. Though cell wall polysaccharides can be damaged when pre-treatments are introduced (Allison *et al.*, 2011; Allison, 2011). Saccharification has also been found to be affected

by *O*-acetylation, whereby hydrolytic enzymes appear to be sterically hindered by acetyl groups (EPOBIO, 2006; B. Yang & Wyman, 2006).

Commercial preparations of enzymes contain more than 80 proteins derived from filamentous fungi *Trichoderma reesei* (Banerjee, Car, *et al.*, 2010b) though commercial cellulose mixtures can contain between 80 and 200 proteins (Banerjee, Car, *et al.*, 2010a). However, the cost of such enzymes currently impedes the development of a practical lignocellulosic ethanol industry (Banerjee, Scott-Craig, *et al.*, 2010). Many researchers isolate proteins from *T. reesie*, such as cellobiohydrolase 2, endo-1,4-β–glucanase 1, β-glucosidase, endo-1,4-β–xylanase 3, β-xylosidase and express them in other organisms (e.g. in yeast *Pichia pastoris*). Alternatively industrial enzymes are used for comparison optimisation of cocktails to release fermentable sugars from the biomass (Banerjee, Car, *et al.*, 2010a).

Table 1.2 Hydrolytic enzymes that can be used to produce platform chemicals in biorefining.

Enzyme family	Enzyme	Substrate	Catalytic activity	Reference
Cellulases	Endo- and exo-cellulases	Cellulose	Break down cellulose chains from within and from each end	(Coombs, 1996; EPOBIO, 2006; King *et al.*, 2009)
	Endoglucanase	Cellulose microfibrils	Release glucose chains from microfibril surface	
	Cellobiohydrolase	1, 4 β-D-Linkages	Releases cellobiose (glucose dimer) from free cellulose chains	
	β-glucosidase	Cellobiose	Hydrolyse cellobiose to glucose molecules	
Hemicellulases	Xylanase	Hemicellulose – xylan, polymeric substrates	Converts polymeric substances into xylooligosaccharides and xylose. Hydrolysis of xylan enables lignin release from paper pulp.	(EPOBIO, 2006; Gao *et al.*, 2011; Subramaniyan & Prema, 2002; Tenkanen, Makkonen, Perttula, Viikari, & Teleman, 1997)
	Mannanase	Hemicellulose – mannan type polysaccharides	Hydrolysis to form oligomers including B-D-manno-oligsaccharides and a range of mixed oligosaccharides containing mannose, glucose and galactose	
	β-xylosidase	Hemicellulose	Degrades xylose chains	

1.3.3 Hemicellulose and hemicellulases

Hemicelluloses are polysaccharides that bind to cellulose and lignin in the plant cell wall to form a strong, resilient network (Cosgrove, 2005; La Grange, Pretorius, & Claeyssens, 2001; Rajoka, Bashir, & Malik, 1997) and can be considered the second most abundant biologically renewable bioconversion resource (Y. Han *et al.*, 2012; Rajoka *et al.*, 1997). They have an array of mass, size and shape characteristics, with the degree of polymerisation ranging from 80 to 200 monomers being insoluble at low temperatures in water, but as temperature increases so does its solubility (Harmsen *et al.*, 2010; ThermoWood, 2003; Walker, Butterfield, Langrish, Harris, & Uprichard, 1993). Esterified and branched polysaccharides xylose, mannose, arabinose, galactose, glucose are linked by glycoside bonds and along with aribino-xylan, gluco-mannan and galactan are all grouped under the family heading of "hemicellulose" thus compositions differ depending on the source (Harmsen *et al.*, 2010; Yarbrough, Himmel, & Ding, 2009). In *Miscanthus* xylose, arabinose and furfural make up hemicellulose (Dee & Bell, 2011). The most common hemicellulose is xylan, which is made up primarily of the five carbon monomer xylose (Harmsen *et al.*, 2010) and is a heteropolysaccharide with arabinosyl, glucuronyl and acetyl residue branches (Ghatora, Chadha, Badhan, Saini, & Bhat, 2006). The xylose monomers bind with β-1, 4 linkages producing xylopyranosyl units forming a sugar backbone. Substituently attached are anhydroxylose units, α-(4-O)-methyl-D-glucuronopyranosyl units and acetyl groups, forming a highly branched structure lacking the crystallinity associated with cellulose (Cosgrove, 2005; Harmsen *et al.*, 2010; Subramaniyan & Prema, 2002). However, xylan structure differs depending on source, extraction method and substitution frequency and type (Ghatora *et al.*, 2006).

Xylanases are a group of hydrolytic hemicellulolytic enzymes that break down xylans to produce xylose (Collins, Gerday, & Feller, 2005; GMO Compass, 2010). As shown in Figure 1.5 the xylan backbone is hydrolysed by endoxylanases (EC 3.2.1.8), β-xylosidases (EC 3.2.1.37) and various debranching enzymes (Ghatora *et al.*, 2006; Puls, 1997). These enzymes have been reported in a range of organisms such as bacteria, algea, fungi, protozoa, gastropods and anthropods (Collins *et al.*, 2005). Exo-xylanases play an important role in releasing monomeric xylose from the reducing end of the xylan backbone. This enzyme type is generally produced by bacteria though endo-xylanases from *T. reesie* also exhibit exo-activity (Juturu & Wu, 2014). Xylanases are produced commercially in genetically optimised *Trichoderma spp*., *Aspergillus spp*., as well as *Bacillus spp*. and *Streptomyces spp.* and can also be found in *Phanerochaetes spp*., *Chytridiomycetes spp*., *Ruminococci spp*., *Fibrobacteres spp*., *Clostridia spp*. (Collins *et al.*, 2005; GMO Compass, 2010; Subramaniyan & Prema, 2002). The removal of lignin is aided by esterases (acetyl xylan esterase, feruloyl esterase, p-coumaroyl esterase) which in turn break bonds between xylose and arabinose side chains, and acetic acid, ferulic acid and *p*-coumaric acid respectively (Subramaniyan & Prema, 2002).

Endo-xylanase converts polymeric substances into xylooligosaccharides and xylose via depolymerisation. Endo-1,4-β-xylanases randomly hydrolyse the xylan backbone, whereas 1,4-β-D-xylosidases break of small oligosaccharides, and side groups are released by α-L-arabinofuranosidase, α-D-glucuronidase, galactosidase and acetyl xylan esterase (Ruller, Rosa, Faca, Greene, & Ward, 2006; Subramaniyan & Prema, 2002; Wong, Tan, & Saddler, 1988). The removal of consecutive D-xylose residues from

the non-reducing end of 1, 4-β-D-xylo-oligosaccharides via hydrolysis is catalysed by exo-1, 4-β-D-xylosidase (Honda & Kitaoka, 2004; Subramaniyan & Prema, 2002).

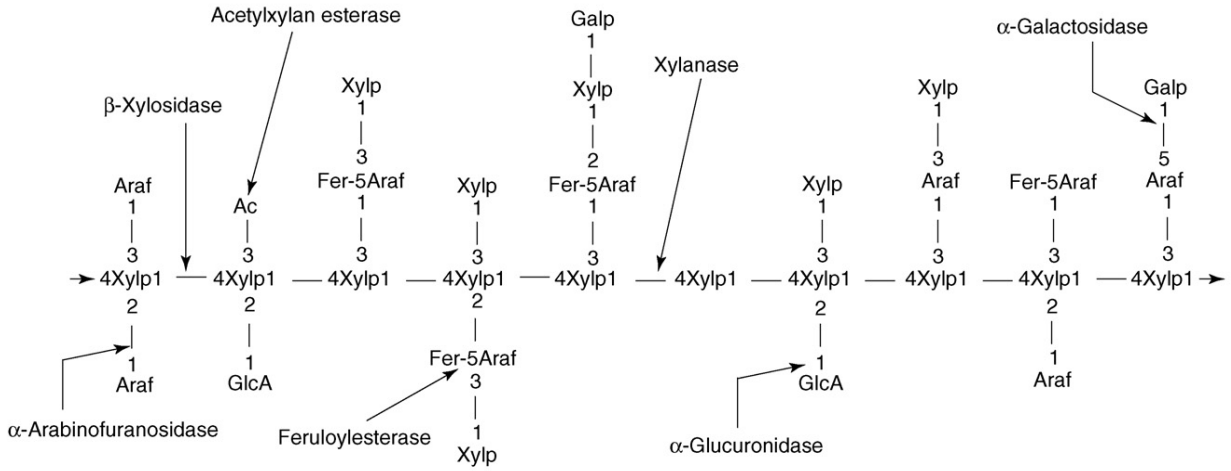


Figure 1.5 Schematic representing the enzymes needed to degrade the polymer corn fibre heteroxylan to monomeric compounds.

Araf - arabinofuranose; Xylp - xylopyranose; Galp - galactopyranose; GlcA - glucuronic acid; Fer - ferulic acid; AC - acetylxylan (Ezeji, Qureshi, & Blaschek, 2007).

1.3.4 Starch and amylases

Though not in the plant cell wall, starch is the most abundant storage carbohydrate in plants (Zeeman, Kossmann, & Smith, 2010). Starch comprises of two glucose polymers, amylose and amylopectin that form the insoluble glucan as shown in Figure 1.6. The life cycle of a plant is dependent upon starch as it is the principal storage carbohydrate. Starch is retained by chloroplasts in photosynthetic organs to then be degraded during the night for respiration and export of sucrose to the rest of the plant (Perez, Palmiano, Baun, & Juliano, 1971; Zeeman *et al.*, 2010). Though *Miscanthus* does produce starch, harvests usually take place when the crop has senesced and therefore would contain little or no starch available for biotechnological uses (Purdy *et al.*, 2013). However, waste streams, such as potato waste could be a starch resource as well as corn (Allen & Pinon, 2013; Esposito & Antonietti, 2015).

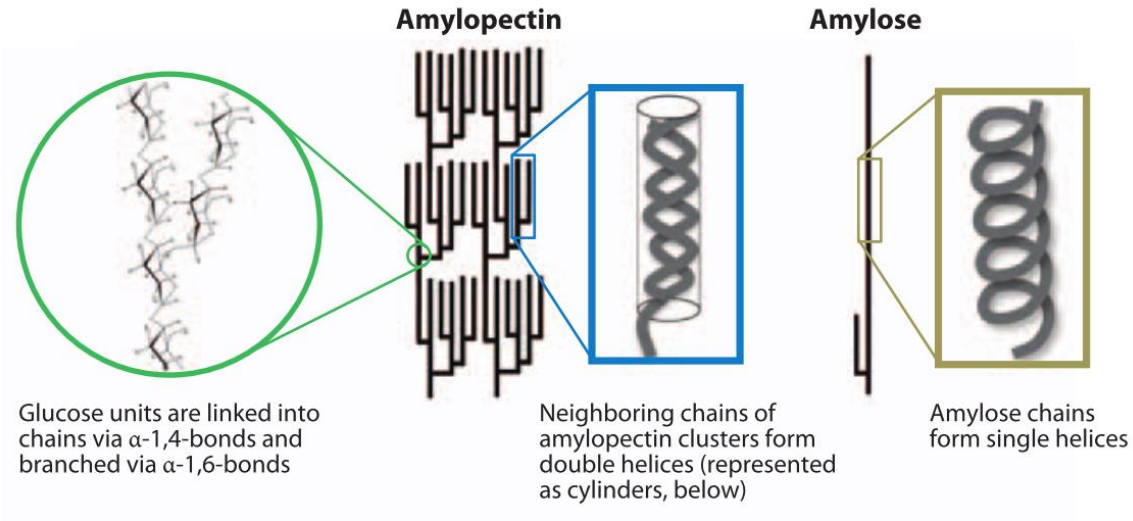


Figure 1.6 Structure of amylopectin and amylose within a starch granule (Zeeman *et al.*, 2010).

Amylases are one group of endo-acting enzymes capable of breaking down the starch molecule via hydrolysis of the internal 1-4 glycosidic bonds (Figure 1.7) resulting in products with alpha anomeric configurations of linear and branched oligosaccharides (Chimata, Chetty, & Suresh, 2011; Regulapati, Malav, & Gummadi, 2007; Xiao, Storms, & Tsang, 2006a). Microbial amylases have become increasingly important in biotechnological investigations (Chimata *et al.*, 2011). Industrial production of amylases is focused on microbial sources due to certain advantages over other sources (such as plants). Advantages include the ease of process modification and optimisation, economically efficiency, consistency, and fewer requirements for production, such as space and time (Sivaramakrishnan, Gangadharan, Madhavan, Soccol, & Pandey, 2006). Amylases have massive a diversity of function throughout a range of industrial applications, making them an interesting group of enzymes to study and use (Chimata *et al.*, 2011). However glucoamylase is also required to hydrolyse the

α-1,6 bonds in amylopectin to release glucose monomers which can then be fermented for the production of ethanol (Borglum, 1980; Chethana *et al.*, 2011).

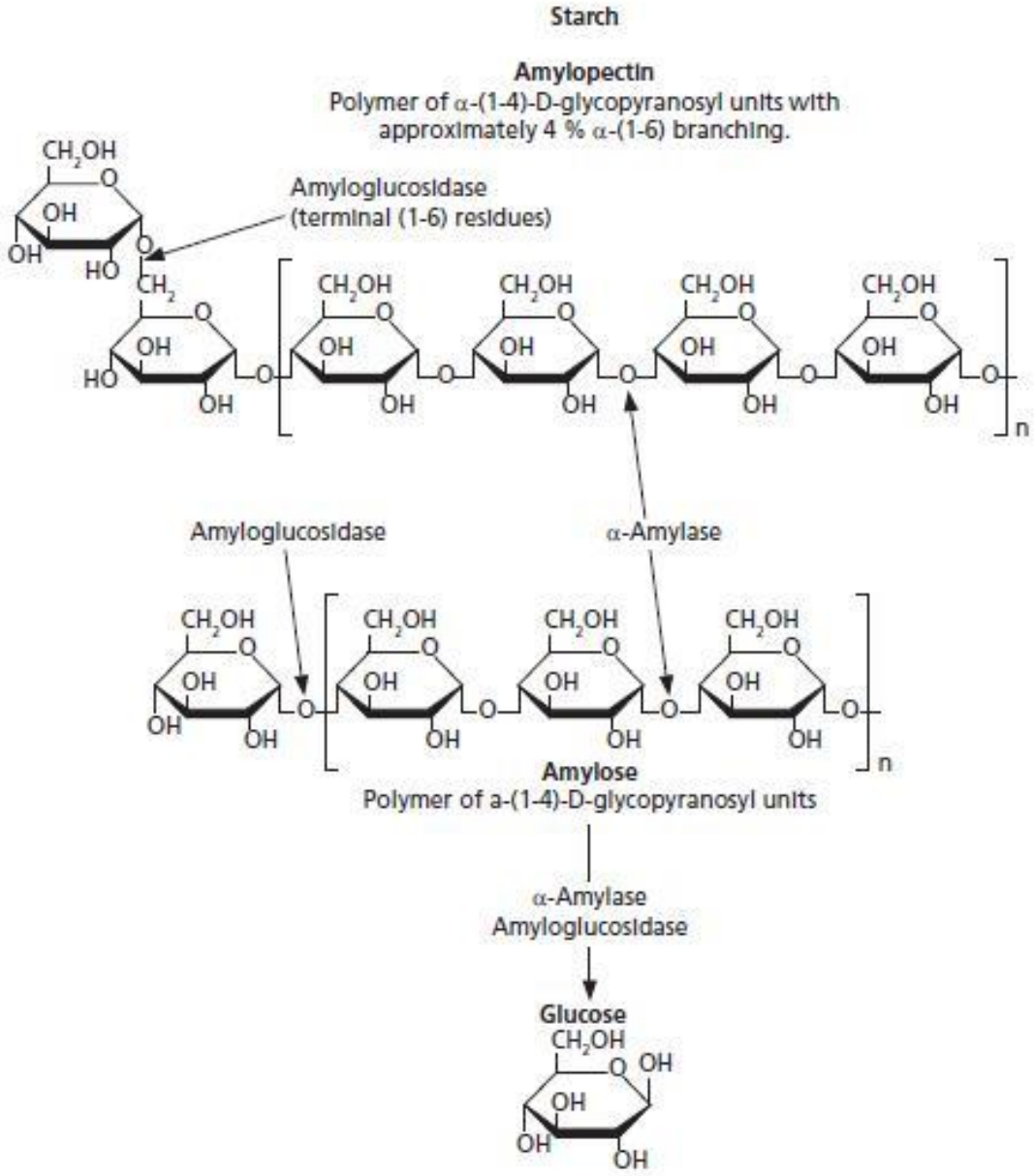


Figure 1.7 Starch molecule degradation by enzymatic action (Siegrist, 2013).

The main feedstock used in biorefineries is still based on starch, involving the fermentation of sugars from sugar and starch crops, with little future for process improvements. Such feedstocks draw from the food stream and therefore compete with food and feed (H. Cheng & Wang, 2013). Alpha-amylase (EC 3.2.1.1) from *Bacillus licheniformis* is used industrially and known as Termamyl by Novozymes. Engineered to be expressed in *Bacillus subtilis,* Termamyl is active up to pH 11 and up to 100 °C and contains a signal peptide and three subunits (Turner, Mamo, & Karlsson, 2007a, 2007b). Termamyl has improved heat stability, reduced dependency on cofactors (calcium) and higher activity at lower temperatures (Novozymes, 2013a). However, Syngenta have engineered an alpha-amylase (section 1.6) protein and incorporated it into the dry grind ethanol process used in industry to produce ethanol from corn (Allen & Pinon, 2013; Enogen Advisory Council, 2011; Halford, 2011; Johnson, Markham, Samyoylov, & Dallmier, 2006; Pinon & Allen, 2013a, 2013b; Syngenta Seeds Inc, 2005; Syngenta, 2013a, 2013b).

1.3.5 Pre-treatment of biomass

Lignin content affects lignocellulosic hydrolysis as it establishes an enzyme proof barrier (Yoshida *et al.*, 2008) as can be seen from Figure 1.3. Converting lignocellulosic material into ethanol can only be achieved when hydrolysis of plant cell wall polysaccharides cellulose and hemicellulose have been turned into fermentable sugar monomers (King *et al.*, 2009). Enzymes, such as laccases, are present in nature that can break down lignin and thus release the cellulosic content for a higher sugar hydrolysis. However, generally, pre-treatment (physical or chemical disruption) is required to increase the hydrolysis of biomass and is used in combination with enzymatic processing (Yoshida *et al.*, 2008) to break down the plant cell wall into

fermentable sugars (Banerjee, Car, Scott-Craig, Hodge, & Walton, 2011; Brandt *et al.*, 2013). The pre-treatment process is able to disrupt the crystalline structure of cellulose, which is highly resistant to depolymerisation, to increase accessibility to the sugars during hydrolysis (H. Cheng & Wang, 2013; Chin & H'ng, 2013; Mosier, Wyman, *et al.*, 2005). Acids or bases are generally used to separate hemicelluloses from the lignin fraction and thus aid fermentation (H. Cheng & Wang, 2013). However, some pre-treatment methods solubilise the lignin fraction and cause inhibition during enzymatic hydrolysis and bioconversion. Many pre-treatment methods use high temperatures and/or acidic conditions to break up the biomass molecular structure so that enzyme hydrolysis can occur (T. H. Kim & Lee, 2005). Pre-treatment of recalcitrant lignocellulosic biomass to initiate its breakdown to simple sugars adds cost to the biorefining process causing a bottleneck (S. J. Lee, Warnick, Leschine, & Hazen, 2012; S. J. Lee, Warnick, Pattathil, Alvelo-maurosa, *et al.*, 2012). Current pre-treatment methods can be classified as mechanical, chemical and biological. Mechanical pre-treatment can involve milling, ultrasonic treatment, and steam explosion (Bussemaker & Zhang, 2013; Chin & H'ng, 2013; Grewell, 2008; Harmsen *et al.*, 2010; Jeon *et al.*, 2013; Le, Julcour-Lebigue, & Delmas, 2013; Pilli *et al.*, 2011; Yun *et al.*, 2013). Chemical pre-treatments include: aqueous fractionation, acid hydrolysis, alkaline hydrolysis, ammonium fibre explosion (AFEX), organosolv, oxidative delignification and ionic liquids. Most chemical pre-treatments involve application of reagents at either low or high pH and therefore required washes and neutralising before materials can be subject to saccharification and fermentation (H. Cheng & Wang, 2013; Harmsen *et al.*, 2010; J. Mielenz, 2009; Mosier, Hendrickson, Ho, Sedlak, & Ladisch, 2005; Walsum *et al.*, 1996; Wan & Li, 2011). High temperatures and pressure are also required, which

means the hydrolysate has to cool before saccharifaction or fermentation occurs (Banerjee, Car, Scott-Craig, Borrusch, & Walton, 2010c; Banerjee *et al.*, 2011; Brandt *et al.*, 2011, 2013; Chin & H'ng, 2013; Crowhurst, Mawdsley, Perez-Arlandis, Salter, & Welton, 2003; L. Dawson & Boopathy, 2008; EPOBIO, 2006; Graenacher, 1934; Hallac *et al.*, 2010; Hallett & Welton, 2011; Harmsen *et al.*, 2010; Plechkova & Seddon, 2008; Y. Sun & Cheng, 2002). However, fermentation, enzyme hydrolysis and microbial growth will all be affected by the presence of solvents in the biomass and need to be removed. To reduce costs in this process the solvents must be recycled which involves draining them from the reactor, evaporation, then condensing; this will also reduce the environmental impact (Harmsen *et al.*, 2010; Y. Sun & Cheng, 2002).

Microbes and enzymes can be used as biological pre-treatments, in many cases the addition of microorganisms such as white, brown and soft rot fungi can be used to break down hemicellulose and lignin polymers, due to their abundant and robust enzymes (Almarsdóttir, 2011; Harmsen *et al.*, 2010; Y. Sun & Cheng, 2002). White rot fungi are the most effective for lignocellulosic pre-treatment, attacking the cellulose and lignin, as do soft rots, whereas brown rots degrade mainly cellulose (Y. Sun & Cheng, 2002). White rot fungi such as *Ceriporia lacerate*, *Stereum hirsutum* and *Polyporus brumalis*, can be used to pre-treat biomass at an optimal temperature for up to eight weeks (Almarsdóttir, 2011; J.-W. Lee *et al.*, 2008). By combining biological pre-treatment with other technologies more sugars may be liberated. Otherwise the development/discovery of novel microbes that can rapidly hydrolyse biomass can advance this long and slow, but low cost process (Brodeur *et al.*, 2011). Most pre-treatments are usually followed by enzymatic pre-treatment. Such processing is

expensive due to costs of enzymes, but is effective for fermentation yields as fractions of hemicellulose and cellulose need to be degraded to mono- or disaccharides (Y. Sun & Cheng, 2002). Such pre-treatment leave no or few toxic hydrolysates and highly concentrated sugars (Almarsdóttir, 2011). Enzymatic hydrolysis utilises bacterial or fungal enzymes to catalyse glycosidic bonds. Enzymes are usually added in cocktails of cellulases, xylanases, mannases and other hemicellulases (Chin & H'ng, 2013). Therefore, by obtaining bacterial isolates or enzymes able to withstand extremes of pH and relatively high temperatures problems from pre-treated material could be eradicated and economical and time costs reduced.

1.4 Microbial sources of lignocellulolytic enzymes

Organisms exist that live on lignocellulose, for example in compost, digestive tracts of ruminants and termites, or rotting wood. Here no pre-treatment is necessary for saccharification because, although it takes a relatively long time, enzymatic mechanisms have evolved naturally that allow them to degrade material (EPOBIO, 2006). Hydrolysis of plant biomass is more complex than that of cellulose. The plant cell wall is a heterogeneous matrix made up of polysaccharides that have varied composition and linkages, as well as aromatic compounds found in lignin (King *et al.*, 2009). It differs considerably between C3 and C4 grasses, and between angiosperms and monocots, especially in the lignin content and structure (Shrestha *et al.*, 2011). Sequencing of microorganisms has shown that more than 50 genes target polysaccharide degradation even in simple bacteria, which often form large communities of thousands of species. Such organisms have highly specialised and diverse enzyme systems (EPOBIO, 2006). It has been shown that fungi, such as *Trichoderma reesei*, that break down grasses, and those that degrade wood, such as

Phanaerochaete chrysosporium, are unlike one another (Ahmad *et al.*, 2011). They are not synonymous in degradation strategy, and therefore may not be optimal for different plants. Consequently cellulosic enzymes from one system may not depolymerise another species (Shrestha *et al.*, 2011). *Clostridium thermocellum* is the most cited microbe for cellulose degradation, though it is anaerobic meaning nitrogen (N_2) or carbon dioxide (CO_2) have to be added to the fermentation rather than air (Roberts, Gowen, Brooks, & Fong, 2010; Rubin, 2008; Tomes *et al.*, 2011). *Cellulomonas uda* may be an aerobic equivalent as well as having hemicellulolytic enzymes (P. Kumar, Barrett, Delwiche, & Stroeve, 2009; Rapp & Wagner, 1986).

Plant pathogens appear to be a novel untapped resource for hydrolytic enzymes. The fungi *Gibberella zeae*, which causes *Fusarium* head blight (González-Fernández, Prats, & Jorrín-Novo, 2010), and *Ustilago maydis*, produce many enzymes that industrially used *Trichoderma reesei* does not, including families of the cell wall degrading enzymes glycosyl hydrolases (Couturier *et al.*, 2012; Martinez *et al.*, 2008). Plant pathogens need to be able to penetrate the plant cell wall quickly to avoid plant defences. *G. zeae* produces a complete set of enzymes for biomass hydrolysis, including endocellulases, exocellulases, β-glucosidases, hemicellulases, and ferulic acid esterases (Paper, Scott-Craig, Adhikari, Cuomo, & Walton, 2007). Phytopathogens *Magnaporthe grisea* and *G. zeae*, have been found to contain more cellulases, hemicellulases, pectinases, carbohydrate binding modules, carbohydrate esterases, and polysaccharide lyases compared to *T. reesei* (King *et al.*, 2009). *Clostridium thermocellum* hydrolyses cellulose and produces ethanol during fermentation, thus eliminating the need for a separate saccharification step (Freier, Mothershed, &

Wiegel, 1988). It also generates a number of industrially useful fermentation products such as; hydrogen, acetic acid, formic acid and butanol (Nakayama *et al.*, 2011; Roberts *et al.*, 2010). Therefore, *C. thermocellum* is considered a candidate for metabolic engineering, whereby energy and carbon flow can be directed towards the required fermentation products (Roberts *et al.*, 2010; Wu & Newcomb, 2009).

Proteins isolated from thermophilic bacteria are known to be intrinsically thermostable, due to adaptation to more extreme environments (C. Lee & Zeikus, 1991). Table 1.3 shows a number of bacteria that have been isolated from a range of diverse ecosystems and have cellulolytic enzymes. A number of these isolates have demonstrated thermotolerance and were isolated from soil (Himmel *et al.*, 2010). Bacteria used in biorefining have been isolated from thermic environments, such as *Thermus thermophilus,* which was first isolated from hot springs in Japan (Koyama, Hoshino, Tomizuka, & Furukawa, 1986; Lönn, Träff-Bjerre, Cordero Otero, van Zyl, & Hahn-Hägerdal, 2003; Oshima & Imahori, 1974). A number of microbes that express hemicellulolytic genes are known human pathogens, such as *Klebsiella pneumonia* and would therefore not be ideal for biorefining (J. Liu, Sun, Liu, Nie, & Zhu, 2010). *Clostridium thermosulphurigenes* (replaced by *Thermoanaerobacterium thermosulfurigenes*) has been shown to harbour xylanase enzymes anchored non-covalently to the cell envelope (Brechtel *et al.*, 1999). The *xylA* protein from this species has also been shown to bind two cobalt ions per subunit, of which it has four (homotetramer) (C. Y. Lee, Bagdasarian, & Zeikus, 1990; UniProt Consortium, 2014). Such metal ions are required for stabilisation and activation of the enzyme at higher temperatures (Epting, Vieille, Zeikus, & Kelly, 2005).

A range of bacterial isolates have xylose degrading abilities by possessing the *xylA* gene (xylose isomerase), such as *B. subtilis*, *K. pneumonia*, *Lactobacillus brevis* and *Streptomyces sp*. (Walfridsson *et al.*, 1996). The *xylA* protein is particularly useful in platform chemical production as it converts xylose directly to xylulose for utilisation by the pentose phosphate pathway, without the need for D-xylose reductase (*XYL1*) and xylitol dehydrogenase (*XYL2*) (Jeffries, 2006). However, endo-xylanase, exo-xylanase and beta-xylosidase enzymes are required for initial depolymerisation of the xylan backbone (Sukumaran, 2009). Such enzymes were investigated in Chapter 5, along with cellulases and amylases.

Table 1.3 Cellulolytic bacteria isolated from diverse natural ecosystems (Himmel *et al.*, 2010).

** - aerobic; [T] – thermophilic; [S] – whole genome has been sequenced; [C] – contigs can be found on NCBI.*

Species	Source	Reference
Acidothermus cellulolyticus[TSC]*	Acidic hot spring	(Koeck, Pechtl, Zverlov, & Schwarz, 2014)
Anaerocellum thermophilum[TSC]	Thermal springs	(Kataeva *et al.*, 2009)
Bacillus brevis[SC]*	Termite gut	(Himmel *et al.*, 2010)
Caldibacillus cellulovorans[T]	Compost	(Bergquist *et al.*, 1999)
Caldicellulosiruptor lactoaceticus[TSC]	Alkaline hot springs	(Mladenovska, Mathrani, & Ahring, 1995)
Caldicellulosiruptor saccharolyticus[TSC]	Thermal pool	(D. Han, Xu, Puranik, & Xu, 2014)
Cellulomonas fimi[SC]*	Soil	(Himmel *et al.*, 2010)
Cellvibrio japonicas[SC]*	Soil	(Himmel *et al.*, 2010)
Clostridium stercorarium[TSC]	Compost	(Koeck *et al.*, 2014)
Clostridium thermocellum[TSC]	Sewage and soil	(Koeck *et al.*, 2014)
Cytophaga hutchinsonii[SC]*	Soil, compost	(Himmel *et al.*, 2010)
Fervidobacterium islandicum[TC]	Host spring	(Koeck *et al.*, 2014)
Paenibacillus polymyxa[SC]*	Compost	(Himmel *et al.*, 2010)
Pseudomonas fluorescens[SC]*	Soil, sludge	(Himmel *et al.*, 2010)
Pseudomonas putida[SC]*	Soil, sludge	(Himmel *et al.*, 2010)
Rhodothermus marinus[TSC]*	Marine springs	(Bergquist *et al.*, 1999)
Saccharophagus degradans[SC]*	Rotting marsh grass	(Himmel *et al.*, 2010)
Sorangium cellulosum[SC]*	Soil	(Himmel *et al.*, 2010)
*Streptomyces rochei**	Termite gut	(Koeck *et al.*, 2014)
Thermobifida fusca[TSC]*	Compost	(Himmel *et al.*, 2010)
Thermotoga maritima[TSC]	Geothermally heated sea floors	(D. Han *et al.*, 2014)
Thermotoga neapolitana[TSC]	Hot spring	(Koeck *et al.*, 2014)
Spirochaeta thermophila[TSC]	Aquatic	(Angelov, Loderer, Pompei, & Lieb, 2011)
Xanthomonas sp.[SC]*	Brack water	(Koeck *et al.*, 2014)

1.4.1 Glycosyl Hydrolases

Proteins from the glycosyl hydrolase family 61 (GH61) have been found to be up-regulated in biomass degrading fungi when grown on cellulose. These proteins may be analogues to carbohydrate binding molecules (family 33) that have been found to utilise metal ions and electron donors to cleave polysaccharides in crystalline formations (Westereng *et al.*, 2011). Carbohydrate binding molecules are often contained within the enzyme, which may help substrate binding and perhaps disrupt the substrate (Fontes & Gilbert, 2010; Himmel *et al.*, 2010). GH61 proteins exhibited in the *Phanerochaete chrysosporium* secretome, along with cellobiose are up-regulated in the presence of xylan, and are able to cleave cellulose synergistically with cellulases using a similar oxidative mechanism (Hu, Arantes, & Saddler, 2011; Westereng *et al.*, 2011). Both plants and microbes possess these enzymes; plants use them to manage cell wall expansion and degradation, energy uptake, and signalling molecule turnover (Davies & Henrissat, 1995). Microbes that possess enzymes that attack similar areas to these will be more effective pathogens, able to reach more protected carbon sources. Glycosyl hydrolases are also able to hydrolyse reactions with lower energy (Dodd & Cann, 2009; Nelson & Cox, 2005). Classification of these enzymes was based on substrate specificity, and now includes protein folding and amino acid sequence similarities due to their direct relationship (CAZy, 2012). At the moment it is possible to have the same enzyme in two or more families because they hydrolyze different substrates. Alternatively different families contain enzymes with differing substrate specificities, which could show evolutionary divergence (Davies & Henrissat, 1995).

1.4.1.1 Multi-Enzyme Complexes

Multi-enzyme complex supplements are often used in saccharification or fermentation (Huang, Su, Qi, & He, 2011). A whole range of enzymes can be used in synchrony to liberate monosaccharides for downstream use, such as conversion to ethanol. B-glucosidase addition to a cellase treatment removes cellobiose inhibition; with a further addition of xylanases cellose becomes more accessible. Arabinofuranosidase increases access to the xylan backbone. A core set of purified enzymes can be used to create a highly efficient enzyme system. Along side the core set, accessory enzymes can be utilised to create a more efficient multi-enzyme system (Huang *et al.*, 2011; M. Zhang, Su, Qi, & He, 2010). However, to enhance thermotolerance and resistance to inhibitors (such as ethanol, phenolics, acids and furans) are generally necessary for more efficient yields. Tolerance enhancements include genome shuffling, artificial transcription factor engineering and random mutagenesis (Huang *et al.*, 2011).

1.5 Metabolic Engineering and Synthetic Biology

1.5.1 Metabolic Engineering

The alteration of an organism, so that it can deal with conditions and the restrictions of cost effective biorefinery, is known as metabolic engineering. This technique uses recombinant DNA technology to improve production, formation and cellular properties via biochemical modification or introduction of new pathways (Akinterinwa, Khankal, & Cirino, 2008; Sauer, Porro, Mattanovich, & Branduardi, 2008). In order to compete with current petroleum derived products, the chemicals produced by microorganisms must be of high yield, titre and productivity (Jarboe *et al.*, 2010; Stephanopoulos, 2007). Accumulation of waste or by-products from fermentation, or the need to function at different pH or temperature, can have inhibitory effects on the

microbe (F. Zhang, Rodriguez, & Keasling, 2011). Microorganisms with desired traits can thus be engineered to not produce these inhibitors. For example, in the production of citric acid, *Aspergillus niger* has been modified so that it no longer generates oxalic acid (Jarboe *et al.*, 2010; Sauer *et al.*, 2008).

Technology has now evolved, whereby modifications and new designs for metabolic pathways, proteins and sometimes even whole organisms can be developed (Jarboe *et al.*, 2010; Stephanopoulos, 2007). However, changes in metabolic activity, growth rate, productivity and viability, caused by stress when the fermentation is scaled up, can decrease the success rate, because of the microbe's need to maintain its natural equilibrium (Jarboe *et al.*, 2010; Sauer *et al.*, 2008). To be industrially viable, the microbe therefore must be able to withstand the production environment. It has been found that reactive oxygen species are key to cellular stress, as they damage the cells in lactic acid production via *Saccharomyces cerevisiae*. By engineering the yeast to generate ascorbic acid, reactive oxygen species levels decrease and viability improves, thus production continues. Therefore, the best way for industry to proceed is to exploit natural biodiversity with metabolic engineering to improve production. Focuses on cellular processes such as transportation, and properties, like morphology, could increase performance and reduce costs (Sauer *et al.*, 2008).

Clostridium acetobutylicum naturally produces acetone, butanol and ethanol and was therefore exploited for industrial fermentation (Lütke-Eversloh & Bahl, 2011). Recently, it has been engineered for increased butanol production, via acetone suppression (J. Lee *et al.*, 2012). *Clostridium tryobutyricum* has also been engineered to utilise both glucose and xylose for butanol production by co-expression of xylose

proton-symporter, xylose isomerase and xylulokinase proteins from *C. acetobutylicum* (L. Yu, Xu, Tang, & Yang, 2015).

1.5.2 Synthetic Biology

Similarly to metabolic engineering, synthetic biology aims to compete and eventually phase out petroleum refinery, for generation of sustainable products such as biofuels (French, 2009). Synthetic biology exploits biological functions that do not naturally exist. It involves design and construction of new biological components, like enzymes, or redesign of existing functions (Colin, Rodríguez, & Cristóbal, 2011; Connor & Atsumi, 2010). Cell free synthesis, tissue and plant engineering, and drug discovery are methods of use. It can also be utilised in the modification of microbes for chemical production (Jarboe *et al.*, 2010; Medema, Breitling, Bovenberg, & Takano, 2011). A metabolic pathway can be designed for a particular purpose; for example to generate a specific chemical or biofuel. An organism that exhibits similar properties, and is well characterised, is chosen to be host to this pathway, and the pathway is altered to fit the organism (Colin *et al.*, 2011; Jarboe *et al.*, 2010). For the new pathway to work, accurately and as expected in the organism, systems biology and metabolic evolution are used (Jantama *et al.*, 2008; Xueli Zhang *et al.*, 2009). This technology has been demonstrated, including the production of chemicals and fuels from biomass, including ethanol, butanol, succinate, malate, xylitol and D- and L-lactate (Jantama *et al.*, 2008; Jarboe *et al.*, 2010; Xueli Zhang *et al.*, 2009). Alternatively, synthetic biology could be used to develop a new organism that would fulfil the biomass conversion criteria (French, 2009).

1.6 Commercial development of biorefining

Biorefining enables high value bio-based chemicals and materials to be converted from building blocks; natural or intermediate products created by microbes via metabolism during fermentation (Esposito & Antonietti, 2015). Lactic acid can be converted into polylactides, these can be used instead of polyethylene derived from petroleum refinery (Sauer *et al.*, 2008; Villegas & Gnansounou, 2008). These molecules contain multiple functional groups, formed by the saccahrification of sugars found in biomass, (Figure 1.2). New families of useful molecules can be obtained from the transformation of these building block chemicals. Sugars are converted into building blocks via chemical or biological processes. Building blocks are transformed into secondary chemicals or derivative families, through chemical conversion (Esposito & Antonietti, 2015). The secondary chemicals were not shown in Figure 1.2, as this project aims to establish microorganisms and enzymes that are able to break down feedstocks, to release sugars that could be processed into products. It is not only the sugars from plant feedstocks that can be used however (Figure 1.2), lignin can be degraded into gallic acid and ferulic acid, which can be converted into secondary chemicals and intermediates. Speciality chemical intermediates can be used in safe food supply in packaging and preservation. Polycarbonates can be used in housing, recreation and health and hygiene industries (Werpy & Petersen, 2004). Production of building blocks influence industrial uses (Figure 1.2). Safe food supply includes pH control agents derived from glutamic acid, and plasticizers derived from fumaric acid. Housing industries could utilise bisphenol A replacements (used in epoxy resin and polycarbonate production) derived from levulinic acid. Polyurethane, which can be used for seals, cradles, tyres button covers is derived from sorbitol. Polypyrrolidones

used for films and fibres can be derived from glutamic and aspartic acid (Crowther, 1956; Polyprog, 2012; Werpy & Petersen, 2004). This project aims to identify microbes and enzymes that can release sugars, which could then be fermented into building blocks or secondary chemicals and intermediates, to aid biorefining and further sustainability, as well as make profit from high-value chemicals.

Candidate chemicals that were not included in Energy Efficiency and Renewable Energy top 12 value added chemicals from biomass were ethanol, ethylene glycol, ethylene oxide, acetone, butanol and fructose. These chemicals are not included because they are already considered super commodities, or have potential to become a super commodity (Werpy & Petersen, 2004). Most of the top 12 chemicals listed can be produced via sugar fermentation with aerobic bacteria, fungi or yeast. However, some chemical and enzymatic transformations rather than fermentation are also important for production of 3-hyroxy butylrolactone, levulinic acid and 2,5 furan dicarboxylic acid. Biological transformations can be adapted by industry to create specific molecular structures. More mild conditions are used than in chemical transformation, at the moment, though extremophiles are being investigated (Ferrer, Golyshina, Beloqui, & Golyshin, 2007; Werpy & Petersen, 2004). Of the top 12 listed value added chemicals from biomass, nine can be produced via bacterial aerobic fermentation: 3-hydroxy propionic acid, glycerol, aspartic acid, fumaric acid, malic acid, succinic acid, glutamic acid (currently a commercial product), glucaric acid, and sorbitol (Werpy & Petersen, 2004).

It is not just the chemicals that can be produced from fermentation that are important. Applications of xylanases include the improvement of bakery products and

animal feed biotechnology, along with wine and juice clarification and paper bleaching. The latter requires thermophilic, alkalophilic enzymes and omits the use of chlorine for a more ecologically friendly process (Ghatora *et al.*, 2006). Novozymes and Terranol released an engineered yeast able to break down C5 sugars, including xylose and arabinose, to make biorefining and the production of cellulosic bioethanol cheaper and more efficient (Novozymes, 2012). Novozyme have also partnered the opening of Beta Renewables' (part of the Mossi Ghisolfi Group) largest advanced biofuels facility in the world, situated in Northern Italy. It has been designed and built specifically for enzymatic conversion of energy crops and agricultural residues to bioethanol (Novozymes, 2013b). A large number of enzymes are commercially available to biorefining industries, such as Acellerase 1000, Novozym 188, Cel7A and Cellic CTec2/HTec2. These enzymes have been isolated from fungal sources such as *T. reesei*, *Apergillus niger* and *T. longibrachiatum* (Banerjee, Car, *et al.*, 2010a, 2010b, 2010c; Cruz *et al.*, 2013; Jabbour, Borrusch, Banerjee, & Walton, 2013; Novozymes Bioenergy, 2010a). Fewer commercially available enzymes have been isolated from bacterial sources, such as *B. subtilis* with optimal pH values of 4-7 and temperatures of 55 °C (Biocatalysts, 2013; Bryant *et al.*, 2011). By screening for specific attributes, this project intended to identify novel bacterial enzymes, which could become commercially viable, due to their characteristics and potentially out-compete available enzymes (section 5.3.5).

There are many sections of the power industry throughout the world that are dedicated or adapting to biomass conversion to fuel or power. Some such as Drax, who supply 7% of the British population with power, burn renewable biomass

alongside coal to produce electricity (Drax Group plc, 2013). Others, such as fuel giant BP have bought up businesses in Brazil to supply ethanol from sugar cane as vehicle fuel (BP plc, 2013; TheBioenergySite News Desk, 2011, 2012). BP and DuPont, have also invested $500 million to build a commercial scale ethanol plant in the UK (Vivergo Fuels) that could produce 420 million litres of ethanol, as well as 500 000 tonnes of animal feed, every year. The plant also has the potential to produce biobutanol in the future (BP plc, 2013). Recently, the British company TMO Renewables have signed a $500 M contract to build 15 factories to hydrolyse household waste for the production of bio-ethanol across the US (Barley, 2010; BBSRC, 2012). TMO have recently risen in the biofuel market, with activity including a 25-year feedstock agreement with Usina Santa Maria Cerquilho (an associate of Copasucar, Brazil's largest sugar cane to ethanol production company) which will supply 400 000 tonnes of bagasse per year for second generation ethanol conversion (TMO Renewables, 2012a, 2012b, 2013). Canadian company Iogen recently sold their industrial enzyme business (Iogen Bio-Products) to Novozymes, to concentrate efforts on a commercial cellulosic ethanol project in Brazil with Raizen (Donovan, 2013). A mature industry has been built up around cellulase enzymes from *Trichoderma longibrachiatum* along with other fungi for commercial production for biorefining (Ingram *et al.*, 1998).

Syngenta have generated an alpha-amylase that can withstand the dry grind process of corn for ethanol production. By genetically modifying the corn plant (Enogen) via *Agrobacterium tumefaciens* the alpha-amylase gene is expressed during the dry grind process (S.-H. Lee, Kim, & Yi, 2009). Therefore, the need to add liquid alpha-amylase, ammonia or acid are eliminated, as is the need to steam between the slurry tank and

liquefaction (Syngenta, 2013a). The enzyme is chimeric and derived from three alpha-amylase subunits from archaeal order *Thermococcales*. The phosphomannos isomerase gene from *E. coli* was used as a selectable marker during transformation (Bonnette, 2013; Johnson *et al.*, 2006). This technology does already exists and the AMY797E amylase shares 93% homology with BD5088 from Innovase LLC (Syngenta Seeds Inc, 2005; Tarantino, 2013). *A. tumefaciens* was used to transform immature corn embryos of corn event 3272, using a vector carrying T-DNA containing the two enzymes. The alpha-amylase is directed by the corn gamma-zein promoter to specifically degrade the corn kernel endosperm tissue (Novel Foods Section, 2013). Not only is this a high throughput process, but there is an increased ethanol yield (Halford, 2011) and the solid fraction is enhanced (Syngenta, 2013a). However, this model is only built for corn, currently there is no equivalent for non-food biomass crops such as *Miscanthus*.

1.7 Aims and objectives of the project

The need for biofuels and platform chemicals from sustainable resources is clear and more efficient microbes are required to depolymerise biomass. The breakdown of lignocellulosic biomass, such as *Miscanthus*, has not been fully optimised. Pre-treatment methods, such as alkaline hydrolysis, organosolv, and steam explosion are currently used to break apart cell wall components and remove lignin to release sugars. However, white rot fungi can naturally perform this task. Such fungal biological pre-treatments are time consuming (Shrestha *et al.*, 2011). Biorefining industries require a more economic process. Thus single or multiple bacteria, which could work together under the same conditions, could be utilised to release the monomeric sugars from pre-treated biomass. Removing the time constrains of fungal pre-

treatment, and costs of chemical treatments. However, such prokaryotes could be utilised simultaneously or after chemical or fungal pre-treatments. To maximise the amount of sugar released from pre-treated biomass, microbes expressing novel enzymes, such as beta-xylosidases and endo-1,4-beta-xylanases, could be utilised. Optimisation of fermentation strategies, such as acid or alkaline tolerant microbes (or engineered microbes), to fit more efficient fermentation strategies increases cost effectiveness and productivity.

The aim of this project was to isolate bacteria capable of saccharification of plant cell wall components, as a precursor to biorefining from pre-treated biomass. Thus, identifying saccharolytic microbes and enzymes that were able to degrade steam exploded or AFEX treated biomass (such as *Miscanthus*). Whereby monomeric sugars could be released for fermentation into platform chemicals and biofuels, as an alternative to petroleum refinery. The process would be near carbon neutral so that industry would be able to operate on a greener level and help combat climate change. The use of *Miscanthus* as biomass, fewer wash steps and faster process time would all aid carbon neutrality for an industrial process. Therefore the first objective was to isolate bacteria that were able to degrade plant cell wall components (starch, cellulose and xylan), that would be accessable after pre-treatment (Chapter 3). Further objectives included, the characterisation of such bacteria that would then lead to enzyme isolation and expression (Chapter 4), as well as genomic characterisation for bioinformatic analysis (Chapter 5). Whole genome sequencing allowed an insight into the mechanics of saccharification, through pathway analysis and protein modelling. Figure 1.8 demonstrates how the project proceeded to address aims and objectives.

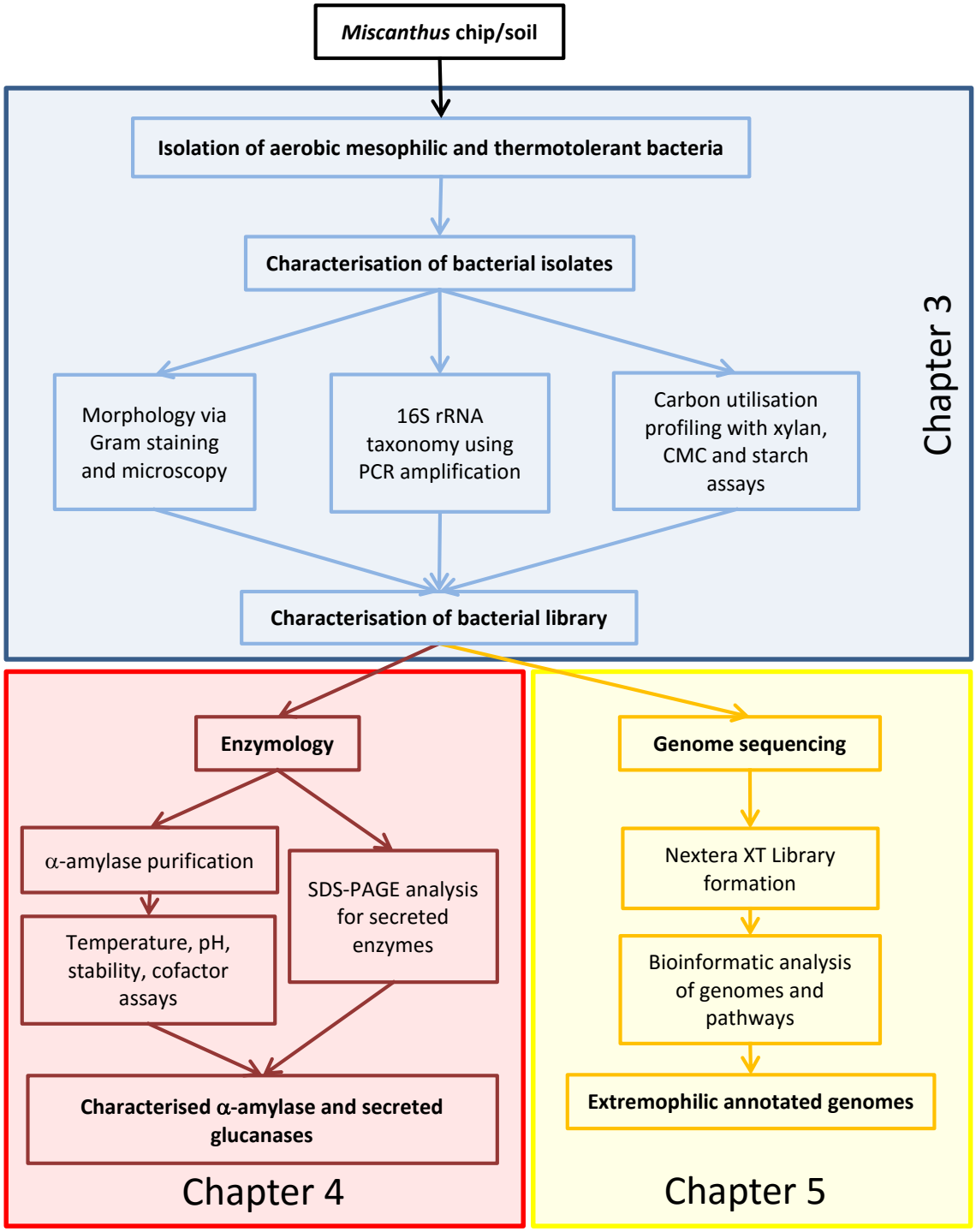


Figure 1.8 The project process with *Miscanthus* chip and *Miscanthus* plot soil as the starting material from which bacteria were isolated.

Blue shows aspects covered in Results chapter 1 (section 3.2); red shows processes covered in Results chapter 2 (section 4.2); yellow shows areas of the project covered in Results chapter 3 (section 5.2).

Chapter 2 Materials and Methods

The experiments contained within this thesis required a large range of materials and methods. Here they are presented in accordance with each results chapter's content, following the logical flow of experiments. Sections 2.2.2, 2.2.3 and 2.2.4 were utilised in Chapter 3; sections 2.2.5 and 2.2.6 were used in Chapter 4, and section 2.2.7 contains methods used in Chapter 5.

2.1 Materials

2.1.1 Kits

BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Life Technologies, Paisley, 4337455)

DNeasy Blood and Tissue DNA Isolation Kit (Qiagen, Redwood City, 69581)

DNeasy® Plant Mini Kit (Qiagen, Redwood City, 69104)

Fast DNA Spin Kit for Soil (MP Biomedicals, Santa Ana, 6560-200)

Gram stain kit (Sigma-Aldrich, Dorset, 77730-1KT-F)

Illustra Exostar™ 1-Step PCR Purification Kit (GE Healthcare Life Sciences, Little Chalfont, US77705)

Meta-G-Nome™ DNA Isolation Kit (Epicentre, Madison, MGN0910)

Nextera®XT DNA Sample Preparation Kit (Illumina, San Diego, FC-131-1024)

Phadebas amylase test (Magle Life Sciences, Lund, 1301)

PureLink™ PCR Purification Kit (Invitrogen, Paisley, K3100-01)

QIAprep Spin Miniprep Kit and a Microcentrifuge (Qiagen, Redwood City, 27104)

QIAquick Gel Extraction Kit (Qiagen, Redwood City, 28706)

Qubit dsDNA High Sensitivity Assay Kit (Life Technologies, Paisley, Q32856)

2.1.2 Materials

1 kb hyperladder (Promega, Madison, G5711)

30% Bis/acrylamide Solution 19:1 (5% C) (Bio-Rad, Hemel Hempstead, 161-0154)

6x blue/orange loading dye (Promega, Madison, G1881)

Acetic Acid (Sigma-Aldrich, Dorset, A6283-500ML)

Acetic acid (VWR, Leighton Buzzard, 20104.334)

Agar High Gel Strength (Melford, Ipswich, M1002)

Agarose (Melford, Ipswich, MB1200)

Ammonium persulfate - APS (Bio-Rad, Hemel Hempstead, 161-0700)

Ammonium sulphate (BDH, Poole, 10033)

Ampicillin sodium salt (Melford, Ipswich, A0104)

Antibody: Anti-His(C-term)-AP (Invitrogen, Paisley, catalog no.: R932-25)

BamHI restriction enzyme (Promega, Madison, R6021)

BCIP/NBT colour development substrate (Promega, Madison, S3771)

Bovine serum albumin (BioRad, Hemel Hempstead, 500-0002)

Broad range molecular weight marker (Promega, Madison, V8491)

Bromo-Thymol Blue (BDH, Poole, 20020)

Butyric Acid (Sigma-Aldrich, Dorset, I5386-5G)

Calcium chloride dihydrate (Sigma-Aldrich, Dorset, 223506)

Calcium chloride dihydrate $CaCl_2.2H_2O$ (Sigma-Aldrich, Dorset, 223506)

Carboxymethylcellulose CMC (Sigma-Aldrich, Dorset, C5678-500G)

Citric Acid (BDH, Poole, 10081)

Coomassie blue R250 (Sigma-Aldrich, Dorset, B-1131)

Crotonic acid (Sigma-Aldrich, Dorset, 113018-500G)

De Man, Rogosa, Sharpe (MRS) Broth (Oxoid, Basingstoke, CM0359)

D-Glucose (Sigma-Aldrich, Dorset, G5767)

diAmmonium tartrate/tartaric acid $C_4H_{12}N_2O_6$ (Alfa Aesar, Ward Hill, 17658)

dNTP's (Promega, Madison, U1240)

EcoR1 restriction enzyme (Promega, Madison, R6011)

Ethanol absolute AnalaR Normapur ACS/R.PE (VWR, Leighton Buzzard, 20821.310)

FastDigest AatII restriction enzyme (Thermo Scientific brand, Life Technologies, Paisley, FD0994)

FastDigest EcoRV restriction enzyme (Thermo Scientific brand, Life Technologies, Paisley, FD0303)

FastDigest SacI restriction enzyme (Thermo Scientific brand, Life Technologies, Paisley, FD1133)

FastStart Taq DNA polymerase (Roche, Burgess Hill, 04738314001)

Formic Acid (Sigma-Aldrich, Dorset, F0507-100ML)

Fructose (Sigma-Aldrich, Dorset, F0127-100G)

Gel Red nucleic acid stain, 10000x in water (Biotium, Hayward, 730-2958)

GeneScan™ 500 LIZ (Applied Biosystems, Life Technologies, Paisley, 4322682)

Glycerol (VWR, Leighton Buzzard, 24387.325)

Glycine (Fisons Scientific Equipment, Loughborough, G/P460/53)

GoTaq DNA Polymerase (Promega, Madison,M3171)

GoTaq®G2 DNA Polymerase (Promega, Madison, M7841)

Hydrochloric acid HCl (Sigma-Aldrich, Dorset, 318949-2L)

Hyperladder IV (Bioline, London, BIO-33053)

IPTG (Melford, Ipswich, MB1008)

Isopropanol/2-Propanol ACS reagent (Sigma-Aldrich, Dorset, 190764)

JM109 Competent Cells, high efficiency (Promega, Madison, L2004)

Kanamycin sulphate (Invitrogen, Paisley, 11815 024)

Lactic Acid (Sigma-Aldrich, Dorset, L1750-10G)

LB Broth (Melford, Ipswich, L1704)

Low range molecular weight marker (BioRad, Hercules, 161-0304)

Lysozyme (Sigma-Aldrich, Dorset, L-6876)

Magnesium sulphate heptahydrate $MgSO_4.7H_2O$ (Sigma-Aldrich, Dorset, 63140 100G F)

Methanol (Fisher Scientific, Loughborough, 10573531)

Micro 1-200 μL pipette tips (VWR, Leighton Buzzard, 613-1080) SDS-PAGE loading

Miracloth (VWR International, Lutterworth, 475855-1)

N,N,N',N'-Tetramethyl-ethylenediamine - Temed (Sigma-Aldrich, Dorset, T-8133)

Ni-NTA agarose (Qiagen, Redwood City, 30210)

Nutrient Broth (Oxoid, Basingstoke, CM0001)

One Shot® BL21 Star(DE3) Chemically Competent *E. coli* (Invitrogen, Paisley, 44-0049)

pET30a(+) DNA (Novagen Merck Millipore, San Diego, 69909-3)

pGEM-T Easy Vector system I (Promega, Madison, A1360)

PhastGel Blue R (Amersham Biosciences, Little Chalfont, 17-0518-01)

POP-7™ Polymer for 3730/3730xl DNA Analyzers (Applied Biosystems, Life Technologies, Paisley, 4335615)

Potassium dihydrogen phosphate KH_2PO_4 (Sigma-Aldrich, Dorset, P8281)

Primers (Biolegio, Nijmegen) (Sigma-Aldrich, Gillingham)

Propionic Acid (Sigma-Aldrich, Dorset, 57400-5G-F)

Protein assay dye reagent concentrate (Bio-Rad, Hemel Hempstead, 500-0006)

Sodium dodecyl sulphate - SDS (Fisher Scientific, Loughborough, S/5200/53)

Sodium Hydroxide NaOH (Fisher Scientific, Loughborough, BPE359-500)

Sodium phosphate dibasic Na_2HPO_4 (BDH, Poole, 10245)

Sodium phosphate monobasic NaH_2PO_4 (Fisher Scientific, Loughborough, 113647)

Sodium sulphate Na_2SO_4 (Sigma-Aldrich, Dorset, 239313-500G)

Starch soluble (Sigma-Aldrich, Dorset, S9765 100G)

Starch soluble ACS (Alfa Aesar, Ward Hill, 36703)

Succinic Acid (Sigma-Aldrich, Dorset, 398055-500G)

Sucrose (Sigma-Aldrich, Dorset, S5391)

TAE (Bio-Rad, Hemel Hempstead, 161-0743EDU)

Tris (Melford, Ipswich, B2005)

Tween 20 (Life Technologies, Paisley, 85114)

Western Blot Transfer solution (BioRad, Hercules, BT0006)

Whatman 0.45 μm filter membrane (GE Healthcare Life Sciences, Little Chalfont, 10401106)

Whatman 1.2 μm filter membrane (GE Healthcare Life Sciences, Little Chalfont, 7191-005)

XbaI restriction enzyme (Promega, Madison, R6181)

X-Gal (Melford, Ipswich, MB1001)

XhoI restriction enzyme (Promega, Madison, R6161)

Xylan from Beechwood (Sigma-Aldrich, Dorset, x4252 100g)

Yeast Extract powder (Melford, Ipswich, Y1333)

Zeba™ Desalting Chromatography Cartridges, 7K MWCO, 10 mL (Life Technologies, Paisley, 89893)

α-Select DH5α Chemically Competent Cells (Bioline, London, BIO-85026)

2.1.3 Equipment

1.5 mL cuvette (VWR, Leighton Buzzard, 634-0676)

1.5 mL sterile tube (VWR, Leighton Buzzard, 700-5239)

14 cm petri dish (Greiner Bio-One GmbH, Stonehouse, 639161)

15 mL Falcon sterile tube (Fisher Scientific, Loughborough, 05-527-90)

2 mL sterile tube (VWR, Leighton Buzzard, 525-0739)

3730 DNA Analyzer - Sanger sequencer (Applied Biosystems, Life Technologies, Paisley, 3730S)

50 mL sterile tube (Fisher Scientific, Loughborough, 14-432-22)

9 cm petri dish (VWR, Leighton Buzzard, 391-0469)

Autoclave (Priorclave Ltd, London, PS/MID/H60)

Bench top orbital mixer (DenLey, Guangzhou Tianhe Longkou)

Bio101 Savant FastPrep cell disrupter (Qbiogene, Inc., Cedex, FP120)

BioPette™ A 4 Pack Plus Liquid Handling Pipetting Package (100-1000 μL pipette; 20-200 μL pipette; 2-20 μL pipette; 0.5-10 μL pipette) (Labnet International, Inc., Edison, P3942-SK4)

Boiling hot water bath (Grant instruments, Cambridge, SBB14)

Class I flow hood (SLEE, London, C17)

Class II Laminar flow biological safety cabinet (NuAire, Plymouth, NU-408FM-400)

Cooled incubator with refrigeration unit (Memmert, Schwabach, IPC400)

Dissection microscope (Olympus, Southend-on-Sea, SZ11)

Dri-Block (Bibby Scientific Ltd, Stone, DB-3)

ExiProgenTM (Bioneer Daedeok-gu, Deajeon, EPG-1203043)

Fine scales (Sartorius Weighing Technology GmbH, Epsom, M-Power AZ124)

Gel electrophoresis rig (Kodak, St. Louis, BioMax Mp1015 and QS710)

Gradient PCR thermocylcer (Somerton Biotechnology Centre, Somerton, GS1)

Hot water bath (Grant instruments, Cambridge, Y6)

HPLC (Jasco, Easton, Column oven: CO-965; Intelligent sampler: AS-1555; Ternary gradient unit: LG-980-02; Plus RI detector: RI-2031; UV/VIS detector: UV-1575; HPLC pump: PU-1580)

IEC Micromax Microcentrifuge (DJB Labcare, Newport Pagnell, 3591)

Incubating mini shaker (VWR, Leighton Buzzard, 444-7083)

Lab Dancer bench top vortex (IKA® Werke GmbH 7 Co. KG, Staufen, 0003365000)

Light box (R.R. Beard, Leighton Buzzard, Illuminator 5000, 4005)

Light microscope (Leica, Milton Keynes, DMLB)

Loop (VWR, Leighton Buzzard, 612-9356)

Magnetic stand (Life Technologies, Paisley, AM10027)

Microscope slides (VWR, Leighton Buzzard, 631-0909)

Mini-Protean Tetra Cell (Bio-Rad, Hemel Hempstead, 165-8000)

MiSeq Desktop Sequencer (Illumina, San Diego, SY-410-1003)

NanoDrop (BioTek Instruments, Inc., Vermont, Epoch™ Microplate Spectrophotometer)

OD_{600} (Biochrom, Cambridge, WPA CO 8000 Cell density meter)

Parafilm M® (Bemis Company, Inc., Brigg, PM996)

pH700 probe (Eutech Instruments Pte Ltd, Landsmeer, ECPH70042GS)

Pipette tips (Starlab, Milton Keynes, 10 μL: S1111-3700; 20 μL: S1110-3700; 200 μL: S1111-0706; 1000 μL: S1111-6701)

Precision standard scales (Ohaus, Parsippany, TS4000D)

Qubit fluorimeter 2.0 (Invitrogen, Paisley, Q32866)

Refrigerated Centrifuge (Eppendorf, Hamburg, 5810R)

Scanlaf Mars Safety Class 2 flow cabinet (Labogene, Lynge, Mars 1500)

Sonicator (MSE, London, Soniprep 150)

Spreader (VWR, Leighton Buzzard, 612-1561)

StarRack™ 96 (Starlab, Milton Keynes, E2396-0013)

StarStore 100 freezer box (Starlab, Milton Keynes, I2310-5846)

Stuart Scientific magnetic stirrer (Bibby Scientific Ltd, Stone, SM1)

Stuart® Magnetic stirrer (Bibby Scientific Ltd, Stone, UC151)

Stuart® Orbital incubator (Bibby Scientific Ltd, Stone, S1500)

Thermal cycler (Applied Biosystems, Life Technologies, Paisley, 2720)

Transilluminator (FluorChem, Alpha Innotech Corp., San Leandro)

Transilluminator (UVP, San Gabriel, Chromato-Vue TM-20)

Wheel tube rotator (LEEC, Nottingham)

2.1.4 Software

3DLigandSite (Structural Bioinformatics Group, Imperial College, London)

AlphaEaseFC™ v4.0.1 (Alpha Innotech Corp., San Leandro)

Artemis.jar Genome Browser and Annotation Tool (Wellcome trust Sanger Institute, Genome Research Limited, Hinxton)

BioCyc (SRI International, Menlo Park, USA)

Biomath calculator (Promega, Madison)

BRENDA (Impressum, Department of Bioinformatics and Biochemistry, Institute for Biochemistry and Biotechnology, Technische Universität Braunschweig)

CAZy (Carbohydrate Active Enzymes, Glycogenomics, Architecture et Fonction des Macromolécules Biologiques, Marseilles, France)

CLC Genomics Workbench v7 (Qiagen, Redwood City)

ClustalW2 2.0 (European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton)

Compute pI/Mw tool (ExPASy, SIB Swiss Institute of Bioinformatics, Lausanne)

EZChrom Elite Clinet/server v3.2 Jasco RI 84 place tray software (Agilent Technologies, Scientific Software Inc., Santa Clara)

ImageJ 1.46r (Wayne Rasband, National Institutes of Health)

KAAS (KEGG Automatic Annotation Server , Kyoto Encyclopedia of Genes and Genomes, Kyoto)

KEGG (Kyoto Encyclopedia of Genes and Genomes, Kyoto)

LAS EZ v2.0.0, (Leica Application Suite, Leica, Milton Keynes)

MEGA v4, v5, v5.1, v5.2, v6 (Molecular Evolutionary Genetics Analysis, McAllister Ave)

MetaCyc (SRI International, Menlo Park, USA)

Microsoft Office (Microsoft UK Headquarters, Reading)

Molecular Toolkit (Colorado State University, Colorado)

NCBI BLAST (National Center for Biotechnology Information, Bethesda)

NCBI ORF Finder (National Center for Biotechnology Information, Bethesda)

NCBI Primer3 (National Center for Biotechnology Information, Bethesda)

Notepad++ (Don Ho, author)

Paint.Net v4.0.5 (dotPDN LLC)

Phyre2 (Protein Homology/analogY Recognition Engine v2.0, Structural Bioinformatics Group, Imperial College, London)

Protein calculator v3.3 (Putnam, 2006)

RAST (National Microbial Pathogen Data Resource, Virginia Bioinformatics Institute)

RDP (Ribosomal Database Project, Michigan State University)

SignalP v3.0 (Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Kemitorvet)

SPECTROstar Omega Spectrophotometer (BMG Labtech GmbH, Ortenberg, S/N 415-1802)

Tm calculator (New England Biolabs, Ipswich, USA)

Transcription and Translation Tool (Attotron Biosensor Corporation, Carson City)

Webcutter 2.0 (Max Heiman, Yale University, New Haven)

2.2 Methods

2.2.1 Sample collection and isolation of bacteria

Bioprospecting for aerobic bacteria took place from both soil surrounding *Miscanthus* plots and composting *Miscanthus* material. Such selection was hypothesised to render microorganisms that had adaptation to their environment, and therefore had potentially novel enzymes. Isolates were then cultured for morphological observations, and gDNA extracted for 16S ribotyping.

2.2.1.1 Isolation of bacteria from soil surrounding Miscanthus plots

Top soil (up to 5 cm depth) was taken from around *Miscanthus* plots. Two grams of soil was added to 50 mL water and heated to 70 °C for 1 hr (pasteurised) (Bahadure, Agnihotri, & Akarte, 2010). One millilitre of solution added to 30 mL De Man, Rogosa, Sharpe (MRS) Broth at three different pHs (4.5, 5.5 and 6.4) in a class II flow hood.

Flasks were incubated at three different temperatures (40 °C, 50 °C and 60 °C) overnight. Serial dilutions were performed up to 10^{-5} and each dilution spread onto MRS agar at pH4.5-6.4, 52 g of ready-made MRS Broth was added to distilled water and made up to 1 L. For solid media 12 g agar was added (1.2% gel strength) and pH was altered with 1 M glacial acetic acid. Autoclaved at 121 °C for 15 min. Plates were incubated across the temperature range until growth appeared (overnight to up to 2 days). Single colonies were obtained and glycerol stocks were made using 500 µL culture and 500 µL 80% glycerol. Stocks were stored at -80 °C.

A loop of glycerol stock was streaked onto MRS agar plates ensuring single colonies. Plates were sealed with parafilm. Plates were incubated at 40 °C overnight. Single colonies were purified onto pre-warmed MRS Agar and incubated overnight at 40 °C, to gain biomass for gDNA extraction. Plates were then stored at 4 °C until used.

2.2.1.2 Isolation of mesophilic bacteria from Miscanthus Chip

Degrading *Miscanthus* chip was taken from the field (Llwyngronw) using nitrile gloves and a zipper sample bag (*Miscanthus* Chip 1). In triplicate, 2 g was weighed out and added to 50 mL of Nutrient Broth in sterile conical flasks whereby 13 g of ready-made Nutrient Broth was added to distilled water and made up to 1 L. For solid media 12 g agar (1.2% gel strength) was added. Both broth and solid media were autoclaved. Conical flasks were mixed at 170 rpm for 15 min. Serial dilutions were then conducted in a class II flow hood. Dilutions 10^{-2} and 10^{-3} were used to isolate bacterial colonies on Nutrient Agar plates. Plates were left to grow at room temperature in a clear plastic box on the lab bench. When colonies had grown they were re-streaked onto new Nutrient Agar plates to gain single colonies. Colonies were cultured in Nutrient broth and glycerol stocks made. Original plates were returned to incubation (unless there

was overgrowth) so slower growing colonies could be obtained and the process repeated.

2.2.1.3 Isolation of thermotolerant bacteria from Miscanthus chip

As before *Miscanthus* chip was sampled from the field (section 2.2.1.2). Nutrient Broth and MRS Broth flasks were agitated at 200 rpm for 45 min at room temperature. Flasks were incubated at 40 °C for 45 min in a water bath, 10 mL aliquots were taken from each flask into 15 mL tubes. Aliquots of 100 μL were plated onto Nutrient Agar or MRS Agar (depending in which media they had been incubated). Flasks were incubated at 50 °C, 60 °C and 70 °C, 10 mL aliquots taken and 100 μL of each aliquot spread on plates. Plates were sealed with parafilm and incubated at constant temperatures (a beaker of water was placed in each incubator to prevent dehydration of media). Stocks were made after being heated to each temperature using 500 μL broth and mixed with 80 % glycerol and stored at -80 °C. Single colonies were then picked and streaked onto new plates and grown at corresponding temperatures to which they had been isolated. Single colonies were then cultured in 1 mL of corresponding broth and glycerol stocks made and stored at -80 °C. The remaining culture was used for Gram staining (section 2.2.3.7). Again the 16S rRNA hypervariabale regions were amplified with a range of primers and sequenced to identify the bacteria (section 2.2.3.1).

2.2.1.3.1 Troubleshooting: Pour plates to Separate Mixed Samples

Molten Nutrient/MRS agar (sections 2.2.1.1 and 2.2.1.2) was allowed to cool to 50 °C then poured into petri dishes in a class II hood. Overnight culture (1 mL) was then added while the agar was still molten and the plate swirled to mix and dilute culture into the agar. Plates were left to solidify, sealed with Parafilm M™ and incubated at the temperature at which colonies were isolated. Colonies were able to grow

throughout the media (if facultative anaerobes) and were separated more than on a spread or streaked plate.

2.2.2 Isolation of bacterial gDNA

Genomic DNA was isolated from bacterial cultures for amplification of 16S rRNA hypervariable regions, which were used for bacterial identification. gDNA extraction was also utilised for isolation of the bacterial chromosome for next generation sequencing. Two different methods were utilised: FastDNA™ Spin Kit for Soil (MP Biomedicals, Santa Ana) and DNeasy Blood and Tissue DNA Isolation Protocol, modification for gram positive bacteria (QIAGEN, Redwood City); following the manufacturer's instructions. DNA was eluted into 50 μL Buffer AE or DES and was then run on a gel to determine size and quality. The DNA was also quantified via NanoDrop.

2.2.3 Bacterial identification

Bacterial isolates were identified using 16S ribotyping via PCR. Troubleshooting of primer efficiency allowed more reliable results to be gained. Isolates were also observed morphologically, using Gram staining and growth curves to determine identity.

2.2.3.1 Polymerase Chain Reaction (PCR) to identify most efficient primer combinations for bacterial identification via 16S rRNA amplification

Eighteen primer combinations (

Table 2.1 2.1 and Figure 2.1) were used to amplify DNA in four bacterial isolates to assess which combination was most effective. The PCR reaction was set up (Table 2.2). Thermal cycler programme: 98 °C for 4 min, followed by twenty-five cycles of: 95 °C

for 30 seconds, 55 °C for 30 seconds, 72 °C for 30 seconds, then 72 °C for 7 min and a hold at 4 °C.

Table 2.1 Primer combination identifier and primer sequences

Primer combination	Approx. length (bp)	Forward primer sequence	Reverse Primer sequence
63f & 1541/20R	1498	CAGGCCTAACACATGCAA	AAGGAGGTGATCCAGCCGCA
63f & 1378R	1324	CAGGCCTAACACATGCAA	CGGTGTGTACAAGGCCCGGGAACG
63f & 1492R	1429	CAGGCCTAACACATGCAA	GGTTACCTTGTTACGACTT
63f & R1401	1338	CAGGCCTAACACATGCAA	CGGTGTGTACAAGACCC
338f & 1541/20R	1203	ACTCCTAGGGGAGGCAG	AAGGAGGTGATCCAGCCGCA
338f & 1378R	1040	ACTCCTAGGGGAGGCAG	CGGTGTGTACAAGGCCCGGGAACG
338f & 1492R	1154	ACTCCTAGGGGAGGCAG	GGTTACCTTGTTACGACTT
338f & R1401	1063	ACTCCTAGGGGAGGCAG	CGGTGTGTACAAGACCC
8f & 1541/20R	1533	AGAGTTTGATCCTGGCTCAG	AAGGAGGTGATCCAGCCGCA
8f & 1378R	1370	AGAGTTTGATCCTGGCTCAG	CGGTGTGTACAAGGCCCGGGAACG
8f & 1492R	1503	AGAGTTTGATCCTGGCTCAG	GGTTACCTTGTTACGACTT
8f & R1401	1393	AGAGTTTGATCCTGGCTCAG	CGGTGTGTACAAGACCC
799f & 1541/20R	742	AACMGGATTAGATACCCKG	AAGGAGGTGATCCAGCCGCA
799f & 1378R	579	AACMGGATTAGATACCCKG	CGGTGTGTACAAGGCCCGGGAACG
799f & 1492R	693	AACMGGATTAGATACCCKG	GGTTACCTTGTTACGACTT
799f & R1401	602	AACMGGATTAGATACCCKG	CGGTGTGTACAAGACCC
8f-AG & 1492R	1501	AGTTTGATCCTGGCTCAG	GGTTACCTTGTTACGACTT
8f-AG & 1378R	1368	AGTTTGATCCTGGCTCAG	CGGTGTGTACAAGGCCCGGGAACG

Table 2.2 PCR master mix for 18 primer combinations.

17 μL of master mix was added to wells plus 1 μL of each primer and 1 μL/1 colony of template.

Reagents	X1 (μL)
dH_2O	14.3
x10 MgCl Buffer	2
dNTPs	0.5
Forward primer (10 μM)	1
Reverse primer (10 μM)	1
Roche Taq (5 U/μL)	0.2
Template	1
Total	*20*

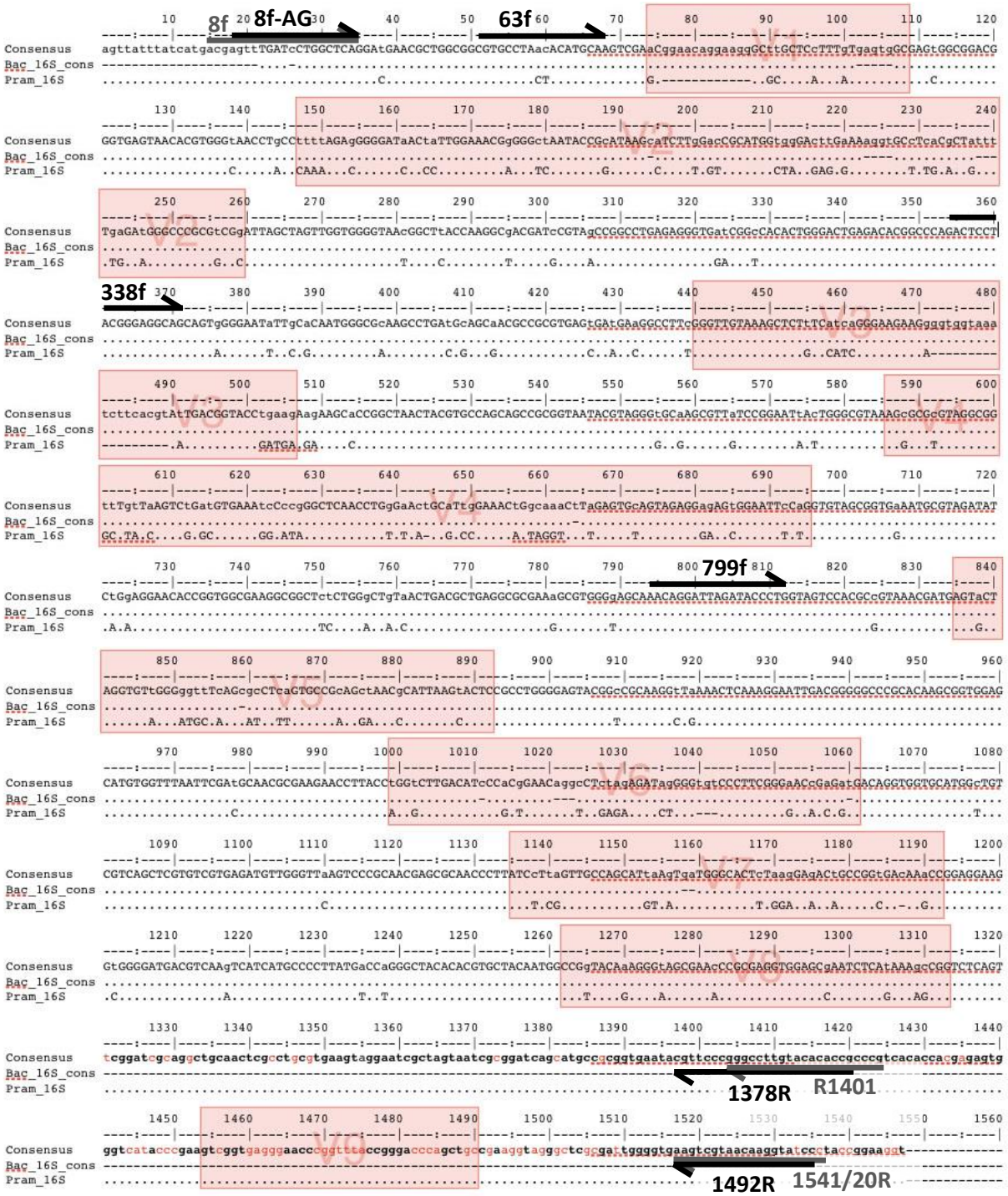


Figure 2.1 Variable regions (V1-V9) of the 16S rRNA gene based on *E. coli* (Brosius, Palmer, Kennedy, & Noller, 1978; The Walser Group, 2012).

Universal primers used for PCR amplification of the 16S rRNA gene of bacterial isolates are imposed on this diagram to highlight which variable regions were amplified with each combination used. Primers 8f/27f have the same sequence (Artursson & Jansson, 2003; Marchesi *et al.*, 1998; Sagaram *et al.*, 2009) and though 1541 has been used extensively (Artursson & Jansson, 2003; Löffler, Sun, Li, & Tiedje, 2000; Marchesi *et al.*, 1998) only an approximation of the sequence could be found on this diagram.

2.2.3.2 Identification of bacteria isolated from Miscanthus via 16S rRNA PCR

Plates with single colonies were stored in the cold room to slow growth. Forward and reverse primers were chosen (8f-AG: AGTTTGATCCTGGCTCAG, and 1492R: GGTTACCTTGTTACGACTT). The master mix was made up for 100 reactions. Single colonies were added to the corresponding wells of the 96 well plate. To select a single colony a 200 μL pipette tip was used to touch the single colony and then mixed in to the master mix in each well. Thermal cycler programme: 98 °C 4 min; 95 °C 30 sec, 55 °C 30 sec, 72 °C 1 min, 72 °C 7mins for 40 cycles, then held at 4 °C overnight. This process was also repeated with 1 μL of culture rather than picking colonies plus a different primer pair: 8f and 1378R (CGGTGTGTACAAGGCCCGGGAACG) using the same master mix recipe (Table 2.2).

Bacterial growth had been streaked onto new plates to gain single colonies. Each plate was given a three letter identifier. 8f-AG and 1492R were chosen as they would produce the largest product covering almost the entire 16S gene and most hyper-variable regions. A number of cultures were mixed colonies and could not be separated. Primer bias was utilised by amplifying the 16S rRNA gene from the same isolate with a range of primer combinations so that both strains could be identified. Gradient PCR was also utilised to determine optimum annealing temperature of primer combinations.

2.2.3.3 Gel electrophoresis and extraction

A 1% agarose gel (Table 2.3) was and loaded with PCR product mixed with 6x blue/orange loading dye (G190A, Promega) on Parafilm M® laboratory. Alternatively, when using GoTaq, green buffer was used so that loading dye did not need to be added. Quantities of PCR product loaded in the gel depended on gel size, as did

voltage and time the gel was run for; 96 well gels: 200 V for 1 hour, other gels were 100-150 V for 40 min. The gel was then visualised using the chromalight on a transilluminator (FluorChem, Alpha Innotech Corp., San Leandro) and a photograph was taken using AlphaEaseFC™ v4.0.1 software (Alpha Innotech Corp., San Leandro).

Table 2.3 Gel electrophoresis recipe for 1% agarose gel.

For 96 well gels reagent quantities were multiplied by four.

Reagent	Amount
TAE buffer	100 mL
Agarose	1 g
Gel Red	5 μL

Gel agarose gel extractions were completed using the QIAquick Gel Extraction Kit (Qiagen, cat no: 28706). Gel extraction took place when more than one band was present per lane on the gel. Bands of the correct size were cut out of the gel in a dark room using a UV box, wearing nitrile gloves (Starlab, Milton Keynes), cuffed lab coat and protective face shield. Qiagen Gel Extraction Protocol (Qiagen, 2001) took place following the manufacturer's instructions. DNA was eluted into 50 μL Buffer EB (10 mM Tris-Cl, pH 8.5).

2.2.3.4 Growth curves

Overnight cultures in MRS Broth were incubated at 50 °C, at 200 rpm. OD_{600} (optical density at 600 nm) was taken the following morning. Three flasks of 30 mL sterile MRS broth were pre-warmed to 50 °C. Overnight culture was used to inoculate the 30 mL sterile MRS Broth (1% in triplicate) and the optical density (OD) immediately taken. Flasks were incubated at 50 °C and 200 rpm, the OD was taken every hour. After 12 hrs of growth the final pH was taken, however the cultures had yet to plateau so

remaining sample left overnight at 50 °C and 200 rpm so that the absorbance and final pH could be taken after 25 hours of incubation.

2.2.3.5 PCR clean-up for improved 16S rRNA gene sequencing

PCR products were cleaned up before sequencing to remove small bands present on the gel or primers still present. Two clean up kits were used to remove any inhibitors and short DNA reads: PureLink® PCR purification kit, and Illustra™ ExoStar™ 1-Step enzymatic PCR and sequence reaction clean up; following the manufacturer's instructions.

2.2.3.6 Sequencing PCR amplicons and analysis

DNA isolated from positive gel bands or cleaned up product (section 2.2.3.5) were sequenced. Forward and reverse primers (1 μL, 10 μM) were added to aliquots of DNA (8 μL). Samples were sequenced via 3730 DNA Analyzer for amplicon analysis. The sequencing reaction consisted of template DNA, primer and Terminator Ready Reaction mix (Table 2.4).

Table 2.4 Master mix for sequencing reaction using 3730 DNA Analyzer for PCR amplicon sequencing.

Reagent	**Quantity (μL)**
Terminator Ready Reaction Mix	4
Template (PCR product)	5-20 ng
Primer	1.6 pM
Deionised water	q.s.
Total volume	*10*

Sequences were analysed using Mega software (v4-6). AB1 files of forward and reverse sequences were viewed as chromatograms to assess their quality. Sequence files were converted to FASTA and trimmed and assembled into the 16S rRNA gene that was covered by the primer pair used. Assembled sequences were aligned to identify those

with high similarity. Sequence files were exported as FASTA files and uploaded to NCBI BLAST and RDP (Ribosomal Database Project) for blastn analysis (Altschul, Gish, Miller, Myers, & Lipman, 1990). Using these two databases, containing microbial 16S rRNA sequences, hits above 97% maximum identity with the lowest e value possible were used to identify isolates. Isolates were also Gram stained (section 2.2.3.7) and studied via dissection microscope to aid 16S identification by comparing results to descriptions in the literature.

2.2.3.7 Gram Staining

Colonies were cultured on agar plates or in 1 mL broth. Slides were labelled using a diamond scriber so that they could not be removed during the staining process. A loop of liquid culture or a loop of distilled water plus a sample of single colony were inoculated onto a slide in a class II hood (three stains per slide). The smear was fixed by passing the slide through a flame three times. Slides were flooded with Crystal Violet solution and left to stand for 60 seconds. The solution was poured off and slides rinsed with tap water and flooded with Gram's Iodine (Lugol's solution) for 60 seconds. The solution was poured off and slides rinsed with tap water. Decolouriser solution was applied to slides until blue dye no longer flowed from each smear and the slide washed with tap water. Slides were counterstained with Safranin solution for 60 seconds then rinsed with tap water. Slides was then blotted with blue roll and left to air dry before viewing under the microscope using the oil emersion lens.

2.2.4 Carbon utilisation of mesophilic and thermotolerant bacterial isolates

Carbon Utilisation Assays (CUAs) were used to characterise enzymatic activity of each bacterial isolate over a range of carbon sources, to assess biorefining capabilities.

CUAs were performed across a pH range to assess acid/alkali capabilities. For thermotolerant isolates a temperature range was also employed.

2.2.4.1 CUAs for mesophilic bacterial isolates

Three assays were performed each with a different carbon source: cellulose, xylan and starch. Carbon sources were integrated into a basal medium (

Table 2.5) and agar added (Bhadra, Rao, Singh, Sarkar, & Shivaji, 2008; Pointing, 1999). pH was adjusted to 4, 5.5 and 6.4 with Hydrochloric acid (HCl); alkaline adjustments (pH 8/8.5, 9/10/11) were made with Sodium sulphate (Na_2SO_4). The media were poured into 14 cm petri dishes with a grid drawn on the back (so that strains could be identified via their 3 letter codes). The colonies were separated into "fast growing" and "slow growing" categories so that one bacterium would not over grow other slower growing samples. Once the media had set a 200 μL sterile pipette tip was used to inoculate each square of the grid by touching the tip into liquid culture (section 2.2.1.2) and then into the solid media. Positive and negative controls (2 μL) were added to wells made in the agar using a cut pipette tip. Plates were sealed with Parafilm M™ and left at room temperature overnight. Plates were then left at 4 °C over the weekend (to slow growth). Growth of single colonies scored and filter paper used to remove growth from the surface of the media. Plates were flooded with 15 mL of iodine (Lugol's solution) until the agar surface was stained and halos around colonies could clearly be seen. Iodine was poured off into an autoclavable container and the surface rinsed with dH_2O. Agar was blotted dry with filter paper and photos taken. The halos were scored on their size and the plates resealed and autoclaved.

Table 2.5 Basal medium recipe plus chosen carbon source for carbon utilisation assays (Pointing, 1999).

Basal medium was heated in a boiling hot water bath before xylan and CMC were added (slowly) on a heated magnetic stirrer. Starch was solubilised in distilled water at 80 °C before pouring into heated basal medium.

g l^{-1} in distilled water	**Reagent**
5.0	diAmmonium tartrate/tartaric acid ($C_4H_{12}N_2O_6$)
1.0	Potassium dihydrogen phosphate (KH_2PO_4)
0.5	Magnesium sulphate heptahydrate ($MgSO_4.7H_2O$)
0.1	Yeast extract
0.001	Calcium chloride dihydrate ($CaCl_2.2H_2O$)
+ 1 % of carbon source	

2.2.4.2 CUAs for thermophilic bacterial isolates

Isolates that were gained from hot temperatures (section 2.2.1.3) were also subject to CUAs described in section 2.2.4.1. The microbes were duplicated at four temperatures (40 °C, 50 °C, 60 °C, 70 °C) on the three different media (CMC, xylan and starch) at pH 6.2. Plates were sealed with Parafilm M™ and incubated overnight at four temperatures (40 °C, 50 °C, 60 °C, 70 °C). Plates were scored for growth, stained with iodine and scored for halos. Isolates with consistent halos over a range of temperatures were highlighted as thermophilic activity is of interest. CUAs were then performed at three additional pHs based on findings in the literature (see

Table 2.6). The pH values chosen were: all carbon sources at pH 4; CMC and starch at pH 8, xylan at pH 8.5; starch at pH 9, CMC pH 10, and xylan at pH 11. Plates were replicated with the bacteria isolated at high temperatures, with each carbon source at each pH at each temperature.

Table 2.6 pH range of known lignocellulolytic enzymatic activity for both purified enzymes and bacterial isolates.

NR – not reported

Enzymatic activity	pH range	Optimal pH	Reference
Cellulase	4-12	NR	(Jüergensen, Ilmberger, & Streit, 2012)
	4-10	NR	(Guevara & Zambrano, 2006)
	6-12	8/9	(Sánchez-Torres, Pérez, & Santamaría, 1996)
	NR	7.2	(Gutierrez-Nava, Herrera-Herrera, Mayorga-Reyes, Salgado, & Ponce-Noyola, 2003)
	6-12	8-9	(van den Burg, 2003)
	NR	5-6	(Zverlov, Mahr, Riedel, & Bronnenmeier, 1998)
Starch degradation	4-9	NR	(Bertoldo & Antranikian, 2002; Niehaus, Bertoldo, Kähler, & Antranikian, 1999)
	4-9	7-8	(Malhotra, Noorwez, & Satyanarayana, 2000)
	4-9	6.5-7	(Gomes, Gomes, & Steiner, 2003)
	NR	5.6	(Melasniewmi, 1987)
	4-6	NR	(Richardson *et al.*, 2002)
Xylanase	4.2-6	5.2	(S. F. Lee, Forsberg, & Gibbins, 1985)
	3.5-9.5	6	(Bronnenmeier, Meissner, Stocker, & Staudenbauer, 1995)
	4-7.5	5-6	(Winterhalter & Liebl, 1995)
	5-8	NR	(McCarthy, Peace, & Broda, 1985)
	3-9	5.4	(Simpson, Haufler, & Daniel, 1991)
	4-12	6-9	(Bansod, Dutta-Choudhary, Srinivasan, & Rele, 1993)
	4-7	NR	(Rajaram & Varma, 1990)
	5-11	6.5	(Khasin, Alchanati, & Shoham, 1993)
	4-11	5-9	(Nakamura *et al.*, 1994)
	5-10.5	8	(Beg, Bhushan, Kapoor, & Hoondal, 2000)
	5-10	NR	(George, Ahmad, & Rao, 2001)

2.2.5 Isolation of enzymes with biorefining potential

With information gained from CUAs, as well as bioinformatics analysis of published genomes, an alpha-amylase gene was selected for analysis. A xylan deacetylase was also selected. The alpha-amylase was later expressed (section 2.2.6) and analysis of other cellulolytic and hemicellulolytic enzymes was carried out using SDS-PAGE and simple plate assays (sections 2.2.6.16 to 2.2.6.18).

2.2.5.1 Isolation of gDNA

The DNeasy Blood and Tissue DNA Isolation with a modification for Gram positive bacteria protocol (Qiagen, Redwood City) was used to extract gDNA from strains (section 0). Candidate genes were targeted for amplification using primers designed based on two genomes available on NCBI. This method was preferred to the FastDNA Spin Kit for Soil and the Meta-G-Nome DNA Isolation Kit as it had a specific modification for Gram positive bacteria.

2.2.5.2 Enzyme gene primer design and reconstitution

Genomes were downloaded from NCBI in GenBank format and opened using Artemis.jar. Genes were found using the "Find/Replace Qualifier Text" function under Edit ("Case Sensitive" was unchecked). A new window opened with a list of genes of this class. Gene(s) of interest were selected and "Bases of Selection" as FASTA for nucleotide sequence or "Amino Acids of Selection" as FASTA for the protein sequence were chosen under "View". NCBI BLAST was used to confirm annotations (blastn for nucleotide, blastp for protein) and for comparisons to similar genes found in other organisms and the GenBank file downloaded. The first 50-60 bases of the amino acid sequence uploaded to SignalP 3.0 (gram positive selected) to identify whether the sequence has a signal peptide plus where it started and ended. The identified signal peptide could then be removed and the primer designed from the base directly after the end of the signal peptide.

The first codon of the enzyme (not including the signal peptide) in the nucleotide sequence was found (if this did not begin with a Methionine (ATG) then the amino acid was added the start). Up to 20 base pairs ending in a G or C were selected and copied

into a Notepad++. Spaces were removed and the sequence capitalised and the primer name inserted. The last 20 bases of the enzyme sequence were selected without the stop codon (TAA, TAG or TGA). The sequence was then reverse complemented in Molecular Toolkit online (Manipulate and Display Sequences) using "inverse complement" and copied back into Notepad++. The complete nucleotide sequence was uploaded to Webcutter 2.0 to identify restriction enzymes that could be used with the vector (in this case any of: EcoRV, XbaI, BamHI, EcoRI, XhoI, with the pET30a C_TERM 6x HIS Polylinker vector). Restriction sites XhoI (CTCGAG) and xbaI (TCTAGA) were then added at the start of each oligo (Table 2.7). The melting temperature of the primers was assessed using Tm calculator (New England Biolabs) to give the melting temperature of each primer and the annealing temperature.

Table 2.7 Enzyme gene primers based on published *B. coagulans* genomes 2-6 and 36D1.

Primer name was based on the start position of the candidate gene within the genomes. *Alpha-amylase AmyA genes were redesigned.

Primer name	Enzyme gene	Sequence	Size of gene (bp)
752	Xylanase	F: TCTAGAATGATGGCATATTTTCCAAATG R: CTCGAGTGCATTTACACTTTGATAGGC	1338
873	Endo-1,4-β-xylanase	F: TCTAGAATGAAAACTTTACCTC R: CTCGAGATTTTTTTCCAAGG	1077
2119	Polysaccharide deacetylase	F: TCTAGAATGCGCGGCAGGAAATTAAATTTC R: CTCGAGTCTTCTCCCCGTTTCAACTGGATG	935
2119NoSP	Polysaccharide deacetylase no signal peptide	F: TCTAGAATGGAAAAGATGTCAGCAAAAGAC R: CTCGAGTCTTCTCCCCGTTTCAACTGGATG	878
753	Xylulokinase	F: TCTAGAATGAAATATGCAATCGG R: CTCGAGAATTTCATCTGTTTTCC	1503
amyF/R	Alpha-amylase*	F: TCTAGAATGGAACGGAATCATACAATCATG R: CTCGAGGTTTTCGCCCGTTTCCTGTAC	1464
723	Endoglucanase	F: TCTAGAATGCCAGTGGCAAAGTTAGATG R: CTCGAGGTCAAAAATAATATGGTTTACTG	1092
Up/down	Up/downstream AmyA	F:CATTCACTTATACTAAACGCATC R: ATGGAACAGTGCCCCTCCTTAG	1533

The primers arrived freeze-dried in 2 mL screw cap tubes. They were centrifuged for 1 min at 14000 rpm. Sterile water (Sigma-Aldrich, Gillingham) was added to the stock primers as specified on the order sheet. New 2 mL screw cap tubes were labelled with corresponding primer names and 450 μL of sterile water was added to each tube. The reconstituted stock was added to the corresponding tube for the working solution (10 μM).

2.2.5.3 Amplification of Enzyme Genes

The enzyme encoding gene primers were tested on three strains of *Bacillus coagulans*. A master mix was made up (Table 2.8) and 45 μL added to each reaction plus 3 μL of 7 different primer pairs (Table 2.7). The thermal cycler was programmed to: 96 °C 3 min; 96 °C 30 sec, 55 °C 30 sec, 72 °C 2 min, 72 °C 7 min for 25 cycles and held at 4 °C overnight. AmyF and AmyR used: 96 °C for 3 min followed by 25 cycles of 96 °C for 30 sec, 60 °C for 30 sec, 72 °C for 2 min, then 72 °C for 7 min and hold at 4 °C. PCR products were then run on a 1 % agarose gel. Gels were run at 100 V for 60 min and photographed (section 2.2.3.3).

Table 2.8 Master mix for enzyme genes amplification from genomic DNA isolated from three strains of *B. coagulans* strains (ABP, ABQ and ABR).

Reagent	**(μL)**
X5 Green GoTaq Buffer	80
dH_2O	280
10 mM dNTPs	8
gDNA (*B. coagulans* ABP)	3
GoTaq Taq Polymerase	4
Total volume	*375*

2.2.5.4 Ligation and Transformation of Alpha-Amylase gene

Amplified alpha-amylase genes were selected for expression in *E. coli* for further characterisation, using pGEM-T Easy Vector System I (Promega, Madison). Following the manufacturer's instructions the alpha-amylase gene was ligated into α-Select DH5α Chemically Competent Cells (Bioline, London). Insert DNA concentration required was calculated using the Biomath calculator (www.promega.com/biomath) to gain 1:3 vector to insert. Ligations were set up (Table 2.9) in sterile 0.5 mL tubes based on NanoDrop results of gene PCR product. The reactions were mixed via pipetting and then incubated for 1 hr at room temperature.

Table 2.9 Ligation mix for alpha-amylase gene ligation into pGEM-T Easy vector.

Reagents	**Standard Reaction (μL)**	**Positive Control (μL)**	**Background Control (μL)**
2x Rapid Ligation Buffer, T4 DNA Ligase	5	5	5
pGEM-T Easy Vector (50 ng)	0.9	0.9	0.9
PCR Product	2	-	-
Control Insert DNA	-	2	-
T4 DNA Ligase (3 Weiss units/μL)	0.9	0.9	0.9
dH_2O	1.2	1.2	3.2
Final volume of	*10*	*10*	*10*

LB agar (Melford) was used for Blue/white screening containing 300 μL of ampicillin, 300 μL of X-Gal (100 mg 5-bromo-4-chloro-3-indolyl-β-D-galactoside dissolved in 2 mL N,N'-dimethyl-formamide and stored at -20 °C) and 300 μL IPTG at 100 μM.

Following the manufaturer's instructions DH5α chemically competent cells were transformed using SOC (Invitrogen Cat. no.: 15544-034). Transformations were then incubated at 37 °C for 1 hr shaking at 200 rpm. After incubation 100 μL was spread

onto LB X/I A “low” plates. The remainder was centrifuged at 7000 rpm for 3 min and 800 μL was disposed of and the pellet resuspended in the remaining 100 μL and plated onto LB X/I A “high” plates. Plates were then incubated at 37 °C overnight so that white colonies could be selected.

Positive transformations (white colonies) were picked from each plate. Colonies were inoculated into 10 mL of LB ampicillin broth (ampicillin 100 μg/mL) and incubated at 35 °C and 190 rpm. Colonies were also inoculated into 200 μL dH_2O for PCR analysis (Table 2.10). The 200 μL inoculated dH_2O was boiled for 10 min then placed on ice. Thermal cycler programme: 94 °C for 2.30 min, followed by 25 cycles of 94 °C for 30 sec, 55 °C for 30 sec, 72 °C for 1.30 min, then 72 °C for 7 min and hold at 4 °C. A 1% agarose gel was run at 115 V for 30 min.

Table 2.10 Colony PCR Master Mix for transformations.

Boiled colony template (1 μL) was added to 24 μL of master mix for each reaction.

Reagent	**X 30 (μL)**
5x Green GoTaq Buffer	120
dNTP’s	15
Primer T7	23
Primer SP6	23
dH_2O	517
GoTaq	10

2.2.5.5 Ligation and transformation of xylan-deacetylase into pGEM-T Easy vector

Ligation and transformation of xylan deacetylase from *B. coagulans* into *E. coli* was approached in much the same way as the alpha-amylase methods (section 2.2.5.4). However the gene would only ligate into pGEM-T and not the expression vector (pET30a) and so the enzyme was not expressed. Troubleshooting included variations of plasmid:insert concentrations as well as using BSA to aid ligation.

2.2.5.6 Plasmid DNA purification using the QIAprep spin miniprep kit and a microcentrifuge of alpha-amylase pGEM-T Easy vector for ligation into pET-30a expression vector

Following the manufacturer's instructions, the plasmid was extracted for transformation into the expression vector. Samples were eluted in 50 μL of EB buffer, then a further 25 μL of EB buffer for an elution total of 75 μL. Plasmids were prepped for sequencing using 3.2 pM of T7 and SP6 primers with 3 μL of plasmid in a 12 μL total volume (section 2.2.3.6).

Alpha-amylase pGEM-T (50 μL) was digested with 12 μL Buffer D, 50 μL dH_2O and 4 μL restriction enzyme (xbaI and xhoI). The reaction was incubated for 2 hr 15 min at 37 °C. The expression vector (pET30a) was digested in the same manner. An agarose gel was run (1 hr at 110 V) to check that both the insert and vector were present. Ligations were set up using the pET30a vector based on NanoDrop results from the gel extraction (section 2.2.3.3) of the insert. Two ligation mixes were used (Table 2.11) for transformation first into Bioline DH5α cells (following the manufacturer's instruction) before BL21 Star DE3 OneShot cells (2.2.6).

Table 2.11 Reaction mixes for ligations of alpha-amylase gene into pET30a vector.

Reaction 1		**Reaction 2**		**Negative control**	
5 μL	X2 Buffer	5 μL	X2 Buffer	5 μL	X2 Buffer
1 μL	Vector	2 μL	Vector	1 μL	Vector
3 μL	Insert	2 μL	Insert	3 μL	dH_2O
1 μL	Ligate	1 μL	Ligate	1 μL	Ligate

Colony PCR was performed for ligations. Colonies were picked and inoculated into 10 mL LB kanamycin broth (incubated at 35 °C at 150 rpm overnight). Plasmids were then extracted in duplicate using the QIAprep Spin Miniprep Kit for transformation into

BL21 Star cells (2.2.6). Colonies were boiled for 10 min then placed on ice for PCR analysis (Table 2.12). Thermal cycler programme: 94 °C for 2 min, then 25 cycles of 94 °C for 30 sec, 50 °C for 30 sec, 68 °C for 1.45 min, then 68 °C for 7 min and hold at 4 °C. A 1% agarose gel was then run at 75 V for 1 hr (2.2.3.3).

Table 2.12 Master mix for *E. coli* colony PCR with pET30a vector.

26 μL of master mix was added to each tube plus 1 μL of template.

Reagent	**X 1 (μL)**
X 10 Buffer	2
dNTP's	0.5
Forward Primer (T7 promoter)	1
Reverse Primer (T7 terminator)	1
dH_2O	21.2
Roche Taq	0.3

2.2.6 Expression of enzymes with biorefining potential

The pET30a expression vector containing *B. coagulans* alpha-amylase gene was transformed into *E. coli* expression cells and induced to produce protein. Invitrogen Champion pET Directional TOPO Expression Kits protocols were followed: Expressing the PCR Product - Transforming BL21 Star™ (DE3) OneShot® cells; Pilot Expression; Protein solubility/insolubility test.

Samples were prepared for the Solubility/Insolubility test with a few differences to the protocol. Pellets were resuspended in 400 μL of Lysis buffer (Table 2.13 and Table 2.14) plus lysozyme and freeze-thawed twice at -80 °C for 5-10 min. Samples were centrifuged for 3 min at maximum speed at 4 °C. Equivalent amounts of supernatant and 4x SDS-PAGE sample buffer (Table 2.19) were mixed. Induced pellets plus the final Non-Induced pellet were mixed with 500 μL of 1x SDS-PAGE sample buffer. 8 μL of sample and 5 μL of protein marker were loaded into the SDS-PAGE gel (2.2.6.1).

Table 2.13 Concentrations of reagents required for lysis buffer used for protein solubility/insolubility test.

Reagent	Concentration
Potassium phosphate, pH 7.8	50 mM
Sodium chloride (NaCl)	400 mM
Potassium chloride (KCl)	100 mM
Glycerol	10%
Triton X-100	0.5%
Imidazole	10 mM
Monopotassium phosphate (KH_2PO_4)	1 M
Dipotassium phosphate (K_2HPO_4)	1 M

Table 2.14 Lysis buffer for protein solubility/insolubility test.

Reagent	Amount
KH_2PO_4	300 µL
K_2HPO_4	4.7 mL
NaCl	2.3 g
KCl	0.75 g
Glycerol	10 mL
Triton X-100	500 µL
Imidazole	68 mg

2.2.6.1 SDS-PAGE

Resolving/running gels were made up (

Table 2.15) poured, approximately 5 mL for each gel. The gel was topped with 400 µL of water saturated butanol and it was left to set for 40 min. The water saturated butanol was washed off with tap water and any residual water was soaked up with filter paper. The stacking gel components were mixed (

Table 2.15), poured on top of the resolving gel with comb and left to set for 30 min.

Table 2.15 Recipe for 2x 12 % SDS-PAGE mini-gels (7 cm).

Acrylamide recipe described in Table 2.20.

Resolving gel		Stacking gel	
Bis/acrylamide	4 mL	Bis/acrylamide	375 µL

Running buffer	2.5 mL	Stacking buffer	625 µL
dH_2O	2.5 mL	dH2O	1.5 mL
10% APS	37.5 µL	10% APS	12.5 µL
Temed	10 µL	Temed	3.75 µL

TGS (tris-glycine-SDS) buffer filled the tank (Table 2.16). Samples were mixed with 4x SDS sample buffer (Table 2.19) and heated to 90 °C for 10 min, then placed on ice. Samples were carefully loaded along with the protein marker (BioRad), using specialised tips (VWR). The gel was run at 175 V for 1 hour 15 min. Coomassie blue (Table 2.21) was used to stain gels with agitation at room temperature overnight. The gel was destained (Table 2.22) until bands could be seen clearly. Destaining was stopped by adding 1 % acetic acid.

Table 2.16 SDS-PAGE Tank buffer. Recipe for a 10x stock solution made up to 1 L.

Reagent	**10x 1 L Tank buffer**
Trizma base	30.3 g
Glycine	144 g
SDS	10 g

Table 2.17 Running/resolving gel buffer.

Dissolved reagents and made up to 250 mL with dH_2O, pH adjusted to 8.6 with HCl. pH was still at 8.6 after 30 min.

Reagent	**250 mL Running Gel Buffer**
Trizma base	45.41 g
SDS	1 g

Table 2.18 Stacking gel buffer.

Reagents were dissolved in 100 mL dH2O and pH adjusted to 6.8 with HCl. pH of trizma base and SDS tested again and adjusted accordingly after 30 min.

Reagent	**100 mL Stacking Gel Buffer**
Trizma base	6.05 g
SDS	0.4 g

Table 2.19 1x SDS-PAGE sample buffer.

Buffer was made up to 10 mL with distilled water; 1 mL aliquots stored at -20 C.

Reagent	1x SDS-PAGE Sample Buffer
0.5 M Tris-HCl, pH 6.8	1.25 mL
100% Glycerol	1 mL
β-mercaptoethanol/DTT	0.2 mL
Bromophenol blue	0.01 g
SDS	0.2 g

Table 2.20 30% Acrylamide.

Made up to 100 mL to dissolve reagents.

Reagent	30% 100 mL Acrylamide
Acrylamide	29.2 g
Bis acrylamide	0.8 g

Table 2.21 Coomassie blue gel stain.

Coomassie blue R250 was dissolved in methanol for 30 min before water and acetic acid were added.

Reagent	1 L Coomassie blue
Coomassie blue R250	1 g
Methanol	500 mL
Dissolve for 30 min on magnetic stirrer	
Distilled water	400 mL
Glacial acetic acid	100 mL

Table 2.22 SDS-PAGE destain.

Reagent	Destain
Ethanol	30%
Glacial acetic acid	10%
Distilled water	60%

The *B. coagulans* alpha-amylase insert was sequenced using the T7 promoter and T7 terminator (section 2.2.3.6) and aligned against previously published genes. The nucleotide sequence was translated using ExPASy (Swiss Institute of Bioinformatics, 2014) and uploaded to Protein Calculator (Putnam, 2006) to gain pI (5.12) and molecular weight (56053.17). The size of the protein (56 kDa) was needed so that it could be identified correctly on the gel.

2.2.6.2 Detecting recombinant fusion proteins via Western blot

SDS-PAGE gels and samples were prepared and loaded as before with 15 μL per sample (section 2.2.6.1). The gel was run for 1 hr 15 min at 175 V. Invitrogen Alkaline Phosphate (AP)-conjugated Antibodies Version B handbook was then followed. The gel sandwich was made and the gel run for 2 hr at 50 V with an ice block in the tank. When the Anti-His(C-term)-AP antibody (5 μL) was added to the dilution buffer the dish was covered in aluminium foil and left agitating overnight at room temperature. Finally the membrane was air dried on filter paper.

2.2.6.3 Expressing the PCR Product - Transforming BL21 StarTM (DE3) OneShot® cells - 1 Litre Pilot Expression

Two 1 L flasks were set up for overnight culture (500 mL LB kanamycin broth in each) of *B. coagulans* ABQ. Overnight culture was added to LB kanamycin broth (1% inoculum), and 2 hr later 100 mL was added to boost OD. Optical densities were taken and recorded. IPTG was added to each flask for induction once OD_{600} 0.5-0.8 had been reached. Aliquots of overnight culture and each time point were centrifuged and the pellet frozen at -20 °C for SDS-PAGE analysis. The remaining culture was centrifuged in 250 mL centrifuge tubes at 9000 rpm for 10 min at 4 °C, pellets were added to 15 mL tubes then frozen at -80 °C.

2.2.6.4 Protein minipreps of 6xHis-tagged proteins from E. coli under native conditions using Ni-NTA

5 mL-10 mL of Lysis buffer (**Error! Reference source not found.**) was added to the acterial pellet. Tubes were sonicated for 15-45 seconds on medium power then placed on ice, and repeated 3-20 times (depending on pellet size). Larger pellets were also freeze thawed twice. Tubes were centrifuged and supernatant transferred to a new

tube. Protocol 14 from "The QIAexpressionist handbook" (Henco, 2003) was then followed. Supernatant was centrifuged at maximum speed at 4 °C. The straw coloured supernatant was then added to chilled 15 mL falcon tubes and frozen at -20 °C. SDS-PAGE gels were set up and run at 175 V for 1 hr and 15 min (section 2.2.6.1). The protocol was continued from Step 7 at room temperature.

2.2.6.5 Protein quantification using Protein Assay Dye Reagent Concentrate (BioRad)

Following the manufacturer's instructions, the dye reagent concentrate was diluted with deionised water. BSA standards were weighed 2-50 mg/mL and 10 mL sample added to 990 mL reagent. Samples were mixed by hand and incubated at room temperature for 5 min. Absorbance was then measured at 600 nm.

2.2.6.6 Enzyme activity assay - Zymography

Native-PAGE was performed with the addition of starch to the gel mixture (Table 2.23) and native tank buffer (Table 2.24) based on a number of protocols found in the literature (Abraham, Nagaraju, & Datta, 1992; J. He, Yu, Zhang, Ding, & Chen, 2009; Upadhyay, Sharma, Pandey, & Rajak, 2005). Starch was solubilised in the water prior to its addition to the gel. After the gel had run it was split, half for zymography, half for coomassie staining. For the zymography gel wash buffer 1 (tris-triton X-100, Table 2.25) was poured into a square dish and incubated at 4 °C for 1 hour. Wash buffer 2 (tris-HCl, Table 2.26) was added and the gel agitated at room temperature for 10 min. The gel was then rinsed with dH_2O and the substrate buffer (Table 2.27) added and incubated with agitation at 50 °C for 1 hour (optimal conditions recorded in the literature), then the substrate buffer changed and rotated overnight at room temperature. The gel was stained with iodine (Table 2.28) for 5 min.

Table 2.23 Two 12% 7 cm SDS-PAGE gel recipe containing 1 % starch for zymography.

Resolving gel		**Stacking gel**	
Bis/acrylamide	4 mL	Bis/acrylamide	375 μL
Running buffer	2.5 mL	Stacking buffer	625 μL
dH_2O	2.5 mL	dH_2O	1.5 mL
Starch	1 mL	10% APS	12.5 μL
10% APS	37.5 μL	Temed	3.75 μL
Temed	10 μL		

Table 2.24 Reservoir buffer of Tris-glycine (500 mL)

As with SDS-PAGE bar the addition of SDS; pH 8.3 (Upadhyay *et al.*, 2005).

Reagent	**Amount**
Glycine	7.207 g
Trizma base	25 mM

Native-PAGE used the same recipe as SDS-PAGE except reducing agents and SDS were omitted. SDS was not added to buffers; tris-glycine was used instead of TGS as the reservoir buffer (Table 2.16); β-mercaptoethanol/DTT (dithiothreitol) was not added to the sample loading buffer (Table 2.19). Samples were not heated before loading onto the gel, and the gel was run with an ice block in the tank to cool the running temperature.

Table 2.25 Zymography Wash buffer 1 (100 mL), pH 7.4 (Upadhyay *et al.*, 2005).

Reagent	**Amount**
Triton X-100	2%
Trizma base	50 mM

Table 2.26 Zymography Wash buffer 2 (100 mL).

Tris buffer adjusted to pH 7.4 with HCl (Upadhyay *et al.*, 2005).

Reagent	**Amount**
Trizma base	50 mM

Table 2.27 Zymography Substrate buffer made in 50 mL 50 mM tris buffer.

Adjusted to pH 5.5 (Upadhyay *et al.*, 2005). *Sodium azide was omitted for final zymogram as it was unnecessary.

Reagent	Amount
Trizma base	50 mM
Calcium chloride ($CaCl_2$)	0.3 g
Sodium chloride (NaCl)	0.01 g
Sodium Azide (NaN_3)*	0.3 g
Triton X-100	0.5 mL

Table 2.28 Lugol's solution for staining starch zymography gel.

Made up to 50 mL in distilled water (Upadhyay *et al.*, 2005).

Reagent	Amount
Iodine (I_2)	0.05 g
Potassium iodide (KI)	0.1 g

2.2.6.7 Enzyme activity assay - starch agar assay

Based on carbon utilisation assays (section 2.2.4.2) and methods in the literature (Mishra & Behera, 2008) purified enzyme solutions (section 2.2.6.4) were loaded into wells made in starch agar (pH 6.4). Samples of induced and non-induced cells were loaded as a positive and negative control of the *E. coli* transformations. Positive control dilutions (100 mg/mL – 0.001 mg/mL of bacterial alpha-amylase, Fisher) and a negative control of boiled alpha-amylase (Fisher) were also loaded. Plates were loaded with 20 μL per sample and incubated at 50 °C overnight. They were then stained with Lugol's solution (Table 2.28).

2.2.6.8 Phadebas amylase test

The test was carried out following the manufacturer's instructions. Modifications included raising the incubation temperature to 50 °C and incubation time extended to 4 hours.

2.2.6.9 TCA protein precipitation

Based on the protocol first described by Sanchez in 2001, supernatant from cultures grown in CMC broth and xylan broth were treated with TCA to precipitate secreted proteins (Sanchez, 2001; Sánchez-Herrera, Ramos-Valdivia, de la Torre, Salgado, & Ponce-Noyola, 2007; Zhao *et al.*, 2011). Protein pellets were mixed with SDS-PAGE sample buffer and run on a 12% SDS-PAGE gel (section 2.2.6.1).

2.2.6.10 Ammonium sulphate protein precipitation and sample preparation

Ammonium sulphate was added to samples to a final volume of 80% saturation (56.1 g/100 mL) (Chandrashekharaiah, Murthy, Narayanaswamy, Murthy, & Swamy, 2013). Samples were mixed with a vortex and centrifuged at 14000 rpm for 20 min. Supernatant was removed to leave the protein pellet intact. Pellets were resuspended in 1 mL 80% ammonium sulphate in dH_2O. The 80% ammonium sulphate was removed after centrifugation at 15000 rpm for 45 min. Pellets were resuspended in 200 μL of 10 mM Tris-HCl (pH 7.5) and mixed via vortex. Solutions were made up to 1 mL with Tris-HCl and 0.561 g ammonium sulphate added to 80% saturation for stable fridge storage.

Samples were prepared for enzymatic assays to prevent ammonium sulphate aggregated. A 100 μL aliquot of resuspended pellet was added to a new tube and centrifuged at 15000 for 15 min. The supernatant was removed and the pellet washed quickly with dH_2O to remove ammonium sulphate, and the supernatant removed. Tris-HCl (20 μL of 10 mM, pH 7.5) was added to washed pellets and mixed via vortex. Distilled water (80 μL) was added to each sample and mixed. This solution was then used for enzyme assays (sections 2.2.6.5 and 2.2.6.7) and SDS-PAGE (section 2.2.6.1).

2.2.6.11 Zeba™ Desalting Chromatography Cartridges, 7K MWCO, 10 mL

Cartridges were used following the manufacturer's instructions whereby 3.5 mL Ni-NTA elutions of *B. coagulans* alpha-amylase protein were used with a 200 µL stacker. Samples were centrifuged for 2 min at 1000 xg.

2.2.6.12 Purified B. coagulans alpha-amylase gene product characterisation - pH range

A pH range was devised to assess the proteins optimum conditions. Thirteen pH buffers described in

Table 2.29, Table 2.30 and Table 2.31 were made up for an extensive range from pH 4.6-10.6 (R. M. C. Dawson, 2002). Starch (1%) was added to a total volume of 10 mL pH buffer. The assay consisted of 40 µL pH buffer plus 40 µL desalted enzyme solution (section 2.2.6.11) per time point. Three time points were used: 30 min (Bischoff, Rooney, Li, Liu, & Hughes, 2006), 2 hours and 4 hours (Puchart *et al.*, 2004; Xiao, Storms, & Tsang, 2006b), to show the optimal pH and time for reaction (J. He *et al.*, 2009). The enzyme/starch buffer solution was heated to 50 °C for 30 min in a thermal cycler. Aliquots (80 µL) were taken plus 20 µL HCl (1 M), and 100 µL iodine added for absorbance (580 nm). pH buffers were doubled in molarity (

Table 2.29-Table 2.31) to prevent substantial pH change. pH altered with temperature so each buffer was heated to 50 °C and the pH tested (Table 2.32).

Table 2.29 McIlvaine buffer solutions pH 4.6-7.

Citric acid monohydrate, $C_6H_8O_7 \cdot H_2O$, molecular weight: 210.14; 0.2 M solution contained 42.028 g/L. Na_2HPO_4, molecular weight: 141.96; 0.4 M solution contained 56.784 g/L. 30 mL 0.2 M citric acid stock solution and 40 mL 0.4 M Na_2HPO_4 were made (R. M. C. Dawson, 2002).

pH	x mL 0.2 M citric acid	y mL 0.4 M Na_2HPO_4
4.6	5.325	4.675

5.0	4.85	5.150
5.6	4.2	5.8
6.0	3.685	6.315
6.6	2.725	7.275
7.0	1.765	8.235
Total	22.55	37.45

Table 2.30 Na_2HPO_4-NaH_2PO_4 buffer solutions pH 7.6 and 8.0.

Na_2HPO_4, molecular weight: 141.96; 0.4 M solution contained 56.784 g/L. $NaH_2PO_4 \cdot 2H_2O$, molecular weight: 156.03; 0.4 M solution contained 64.42 g/L. Solutions were made up to a total of 10 mL per pH (R. M. C. Dawson, 2002).

pH	**x mL 0.4 M Na_2HPO_4**	**y mL 0.4 M NaH_2PO_4**
7.6	4.35	0.65
8.0	4.735	0.265
Total	9.085	0.915

Table 2.31 Glycine-NaOH buffer solutions pH 8.6-10.6.

Glycine (aminoacetic acid), $C_2H_5NO_2$; molecular weight 75.07. 2.5 mL 0.4 M glycine (30.02 g/L) was mixed with x mL 0.4 M NaOH (molecular weight: 39.997) and diluted to 10 mL with y distilled water for each pH (R. M. C. Dawson, 2002).

pH	**x mL 0.4 M NaOH**	**y mL H_2O**
8.6	0.2	7.3
9.0	0.44	7.06
9.6	1.12	6.38
10.0	1.6	5.9
10.6	2.275	5.225
Total	5.635	31.865

Table 2.32 Change in pH with temperature.

Thirteen pH buffers were heated to 50 °C to test the shift in pH with temperature.

pH at 25 °C	**pH at 50 °C**
4.6	4.5
5.0	5.12
5.6	5.52
6.0	6.02
6.6	6.75
7.0	7.18
7.6	7.63
8.0	8.17
8.6	8.78
9.0	9.24
9.6	9.86

10.0	10.45
10.6	11.78

2.2.6.13 Purified B. coagulans alpha-amylase gene product characterisation - Temperature Range

Desalted enzyme solution (40 µL, section 2.2.6.11) and 40 µL of optimal pH buffer were mixed and incubated at each temperature. Temperatures used were: 25 °C (room temperature), 30 °C, 37 °C, 60 °C, 80 °C, 90 °C. Each temperature had seven time points whereby an aliquot was taken, the reaction stopped and iodine added (section 2.2.6.12) so the absorbance (580 nm) could be calculated. These time points were: 10 min, 30 min, 1 hour, 2 hours, 4 hours, 5 hours and 7 hours (Burhan *et al.*, 2003).

2.2.6.14 Purified B. coagulans alpha-amylase gene product characterisation - Thermostability

Stability of the alpha-amylase was assessed at a range of temperatures. Desalted enzymes (section 2.2.6.11) were incubated for 30 min at 50-100 °C (Honda & Kitaoka, 2004) in the optimal pH buffer without starch. Aliquots (80 µL) were added to the optimal pH buffer (80 µL) with starch and incubated at the optimal temperature for 2 hours. The reaction was stopped and 200 µL iodine added so that absorbance (580 nm) could be read.

2.2.6.15 Purified B. coagulans alpha-amylase gene product characterisation - Cofactor Concentration

To test the effect of calcium as a cofactor for the activity of alpha-amylase, a calcium gradient (Table 2.33) was devised (Feller *et al.*, 1992; Ramesh & Lonsane, 1989). Calcium chloride (110.98 g/mol) had a stock concentration of 0.1 M and was added to 80 µL reactions (40 µL desalted enzyme solution plus 40 µL optimal pH buffer with 1%

starch). Solutions were incubated for 4 hours at 40-60 °C. 1 M HCl was added to stop the reaction and 100 µL iodine added so that absorbance (580 nm) could be calculated.

Table 2.33 Calcium chloride gradient to test for optimal cofactor concentration.

X µL stock 0.1 M CaCl was added to 80 µL reactions which were then incubated at the optimal pH, at three temperatures (40, 50 and 60 °C) for 4 hours.

Concentration (mM)	x µL stock 0.1 M CaCl
0	0
2	1.6
3	2.4
4	3.2
5	4.0
6	4.8
7	5.6
8	6.4

2.2.6.16 Xylanase and glucanase discovery

Cultures of selected bacterial isolates were grown in xylan and CMC broth overnight in 50 mL tubes at isolation temperatures (section 2.2.1.3) and 200 rpm. Tubes were centrifuged at 8500 rpm for 30-60 min to pellet cells. Cell free supernatant was dispensed into new tubes. Proteins were precipitated using TCA (section 2.2.6.9) and ammonium sulphate (section 2.2.6.10). Samples were run on SDS-PAGE (section 2.2.6.1), plate assays (section 2.2.6.7), and zymography (section 2.2.6.5) with 1% xylan and CMC.

2.2.6.17 HPLC analysis of xylan and CMC degrading isolates - Jasco sample preparation

Cultures were grown in 30 mL xylan/CMC broth. A 1 mL aliquot was taken from the culture and centrifuged at 14000 rpm for 2 min. The supernatant was then removed to

sterile 2 mL tubes. Supernatant and pellet were stored at -80 °C until use. Aliquots were used in a preliminary investigation of metabolic products via Jasco analysis.

The system was purged before use to remove any air bubbles throughout the equipment. The mobile phase (5 mM H_2SO_4) was degassed to prevent bubble formation. Supernatant was thawed at room temperature and 100 μL added to sterile 2 mL tubes and heated for 10 min at 80 °C. Samples were centrifuged at 14000 rpm for 5 min and 50 μL of supernatant was transferred to new 2 mL tubes. Internal standard (950 μL, 100 mM crotonic acid: 2.15 g of crotonic acid in 250 mL of 5 mM H_2SO_4) was added to the 50 μL sample and mixed. Samples were filtered using a 1 mL syringe and a 0.45 μL filter into glass vials, capped and were loaded into the Jasco. Standards (

Table 2.34) were treated in the same way as samples. The column used was ROA (Rezek).

Table 2.34 Calibration standards for Jasco column RAO (Rezek).

Standard 1	**Dilution**	**Sucrose**	**Glucose**	**Fructose**	**Succinic Acid**	**Lactic Acid**	**Level**
mg/mL		10.69	11.875	11.25	9.745	10.925	
Calibration values	10	0.535	0.594	0.563	0.487	0.546	5
	7.5	0.401	0.445	0.422	0.365	0.410	4
	5	0.267	0.297	0.281	0.244	0.273	3
	2.5	0.134	0.148	0.141	0.122	0.137	2
	1	0.053	0.059	0.056	0.049	0.055	1
Standard 2	**Dilution**	**Formic Acid**	**Acetic Acid**	**Butyric Acid**	**Propionic Acid**	**Ethanol**	
mg/mL		9.89	9.875	9.925	9.895	11.235	

Calibration	10	0.495	0.494	0.496	0.495	0.562	10
values	7.5	0.371	0.370	0.327	0.371	0.421	9
	5	0.247	0.247	0.248	0.247	0.281	8
	2.5	0.124	0.123	0.124	0.124	0.140	7
	1	0.049	0.049	0.050	0.049	0.056	6

2.2.6.18 HPLC analysis of xylan and CMC degrading isolates - Jasco software

EZChrom Elite Clinet/server v3.2 Jasco RI 84 place tray software (Agilent Technologies, 2005) was opened offline. A method was set up via "Sequence" then "Sequence Wizard" and a method was chosen. Sample amount was changed to 0.05 (50 μL sample) and the internal standard set at 1 (950 μL, section 2.2.6.17). The sample was given a unique name and saved to a folder in "Sequence". Injection was changed to 25 μL leaving "First" at 1 and "Increment" at 1. In the Sample List each sample was given a unique name with the first sample injected twice. The "Sample ID" was copied and pasted into "Filename" and a Stop Method (volume: 25 μL) added. Sequence was saved to file. Returning to original window (opened from desktop) the sequence was opened via "File" then "Open" and "Sequence" selected, followed by "Sequence run".

2.2.6.18.1 Result Analysis

The baselines of samples needed to be edited as a negative peak at the start caused the baseline to be lowered. In the open sequence each sample could be selected, then opened, via right click and "Open Data", where the chromatogram could be viewed. "Move Baseline" was used to correct the baseline whereby the incorrect line was selected then dragged up to the correct position. "Split Peaks" was used if one acid appeared close to the next peak so that it appeared as a shoulder. The sequence could then be saved.

2.2.7 Whole genome sequencing of 23 bacterial genomes – Nextera XT library preparation

The whole genomes of 23 selected aerobic bacteria (some facultative anaerobes) were completed, due to their activity on CUAs (section 2.2.4). It was hypothesised that such isolates would encode genes for enzymes that may have adapted to *Miscanthus* degradation, and thus be of use in industrial biorefining.

2.2.7.1 Bacterial gDNA extraction and quantification

gDNA was extracted from selected samples via the FastDNA Spin Kit for Soil (section 2.2.2) starting from liquid culture based on previously reported protocols (Jiang, Alderisio, Singh, Xiao, & Icrobiol, 2005). Samples were grown in 1 mL cultures, then centrifuged at 14000 rpm (2 mins), and 500 μL of supernatant removed. Pellets were resuspended in the remaining supernatant and transferred to Lysing matrix E tubes. Following the manufacturer's instructions FastDNA Spin Kit for Soil standard protocol took place. gDNA was then run on a 1% agarose gel to check quality and Qubit 2.0 Fluorimeter (Invitrogen, Paisley) was used to check quantity. The high sensitivity assay kit (0.2 – 100 ng) was utilised with high sensitivity standards, marked dsDNA HS. Quibit working solution was prepared by combining Qubit reagent (28 μL) and Qubit buffer (5572 μL). Standards were prepared by adding 190 μL of Qubit working solution to 10 μL of each standard and samples were prepared in the same way with 10 μL of each sample.

2.2.7.2 gDNA Clean Up - DNeasy® Plant Mini Kit

To remove any media residues still in gDNA samples the DNeasy kit was used following the manufacturer's instructions from step 13. Samples were eluted into 60 μL DES followed by a further elution in 40 μL for a total of 100 μL gDNA.

2.2.7.3 Nextera XT library preparation

gDNA extracted from the selected cultures (sections 2.2.7.1 and 2.2.7.2) were diluted to 0.2 ng/μL with molecular grade water. Following the manufacturer's instructions, Nextera® XT DNA Sample Preparation Guide (Illumina, Part~ 15031942 Rev. C 2012), libraries were generated. Genomic DNA was fragmented into 300 bp sections and tagmented via transposomes. The indexes were added for primer sequence addition and fragements then amplified. Amplicons were then cleaned up and the library normalised before gDNA being pooled and loaded into the MiSeq cartridge.

2.2.7.4 Bioinformatic analysis - CLC Genomics Workbench

CLC Genomics Workbench 7.03 was used to analyse the raw data produced by Illumina MiSeq (2.2.7). Sequences were imported via the "Illumina" option, selecting the "NCBI/Sanger or Illumina pipeline 1.8 and later" with a minimum read length of 150 bases and a maximum of 600-1000 bases. "Paired end reads" was selected, failed reads removed and files saved in a new folder by bacterial three letter identifier (section 2.2.1). A sequencing report was prepared for each genome and saved in the relevant folders. The default settings were used for "Merge overlapping pairs" and reports saved in corresponding folders. Sequences were trimmed by quality scores using the default settings and report files saved in the corresponding folders. De novo assembly was used to construct contigs which were mapped to reference genomes (if available). Contigs were extracted and saved.

A local BLAST library was used to search for homologous genes within each genome. Candidate genes were found either by downloading available genomes and searching for gene names in artemis.jar or searching NCBI databases. Both protein and nucleotide sequences of genes were copied into Notepad++ and saved. Sequences could then be imported into CLC via the "New sequence" function. Once all protein sequences had been imported a BLAST database was created from extracted contigs from each genome. Multi-blast could then be performed using the contig file against a folder of genes via tblastn so that "Local BLAST database" could be selected. Contigs with homologous genes/conserved regions could then be extracted an opened and the gene sequence found using positions via "show text contents" option. The gene could be saved as a FASTA file and used for modelling.

2.2.7.5 Bioinformatic analysis - RAST SEED Viewer

Rapid Annotation of microbial genomes using Subsystems Technology (RAST) was utilised for genome annotation (Overbeek *et al.*, 2014) using contigs assembled in CLC Genomics (section 2.2.7.4). Contigs were uploaded and RAST searched for homology via gene calling and functional annotation. Results could be downloaded or viewed online. "View details" was selected and the genome could be downloaded, and details of analysis viewed, such as quality checks, warnings, and possible missing genes. View "Close strains for this job" allowed closely related strains to be grouped to the uploaded genome via "Create close strains set". "Browse annotated genome in SEED viewer" gave a pie chart of subsystem category distribution. Features of each subsystem could be viewed and pathways visualised using the "Features in subsystem" tab. "View closest neighbours" gave a list of genomes closest to the uploaded, annotated genome which could be exported. The "Compare" tab allowed the

uploaded genome to be compared to the metabolic reconstruction of other published genomes. Annotations could also be edited using the “Annotate” tab.

2.2.7.6 Protein Modelling - Phyre2

Nucleotide sequences of enzyme genes from homology searches in CLC (section 2.2.7.4) were translated to protein sequences using ExPASy (Swiss Institute of Bioinformatics, 2014) and Transcription and Translation Tool (Attotron Biosensor Corporation, 2014). These protein sequences were then copied into Phyre2 (Kelley & Sternberg, 2009) and a Phyre search completed using both “normal” and “intensive” options. Phyre Inverstigator was used to generated and assessed 3D structural models. Alternatively, protein sequences from annotated genomes in RAST were downloaded in FASTA format. Genes were uploaded to Phyre2 Intensive for modelling.

2.2.7.7 Protein Modelling - 3DLigandSite

Protein sequences were uploaded to 3DLigandSite (Wass, Kelley, & Sternberg, 2010) and ligand analysis run.

2.2.7.8 Pathway modelling – KAAS

Contigs generated in CLC (section 2.2.7.4) were uploaded to KAAS (KEGG Automatic Annotation Server v2.0) via “KAAS job request (BBH method)”. Search programme “GHOSTZ” was selected and 40 relevant bacterial species selected. Assignment method was bi-directional best hit (BBH). Results were compared to proteins found in RAST subsystems and pathways from BioCyc/MetaCyc.

2.2.7.9 Further protein characterisation - Protein Calculator

Protein sequences were uploaded to Protein calculator (Putnam, 2006) to generate the molecular weight, charge over pH range and pI (isoelectric point).

Chapter 3 Isolation and Characterisation of Bacterial Library

3.1 Introduction

As described in section 1.1, a great need for renewable energy and green fuels has been identified. Recalcitrance of the plant cell wall increased processing costs via low digestibility (S. J. Lee, Warnick, Leschine, *et al.*, 2012; S. J. Lee, Warnick, Pattathil, Alvelo-Maurosa, *et al.*, 2012). Isolates were hypothesised to have enzymes for plant cell wall degradation, due to the environment they were isolated from. Bioprospecting for bacterial isolates, containing the enzymes necessary for the deconstruction of the *Miscanthus* cell wall, are best isolated from around the plant, where leaves will have been dropped or from already partially degraded material (Oboh, Ilori, Akinyemi, & Adebusoye, 2006; Piao, Markillie, Culley, Mackie, & Hess, 2013). Here the bacteria have adapted to the environment and feedstock, thus requisite enzymes should be specific and potentially novel (Allgaier *et al.*, 2010; Hendriks & Zeeman, 2009). Isolates could then be of use for biorefining of biomass feedstocks, such as *Miscanthus* (Charlton *et al.*, 2009), via relevant saccharolytic enzymes for monomeric sugar release (Dumon, Song, Bozonnet, Fauré, & O'Donohue, 2012; Lionetti *et al.*, 2010). Therefore, bacteria that exhibited growth were assayed for enzymatic activity, which could be of use after pre-treatment of the biomass. Some isolates with enzyme activity were selected for enzyme isolation (section 4.2) and next generation sequencing (section 5.2). Isolates able to degrade biomass could be used for fermentation to produce ethanol or butanol for biofuels, xylitol for artificial sweetening, along with a number of other products (Stephens *et al.* 2007, Jeffries 2006, Dee & Bell 2011, Vrije *et al.* 2002, Livesey 2003). As discussed in Chapter 1 (1.2) enzymes able to break down

lignocellulose are essential for biorefinery. A number of enzymes classified with cellulolytic or xylanolytic function have also been found to be bifunctional, where by the enzyme can act on both cellulose and xylan (Pérez-Avalos, Sánchez-Herrera, Salgado, & Ponce-Noyola, 2008). Screening the same bacteria on a range of carbon utilisation assays may highlight such enzymes as well as those able to withstand a range of pH and temperatures. Many xylanases have optimal temperatures of 45-60 °C (Biocatalysts Ltd., 2015a, 2015b; Biocatalysts, 2013; Novozymes Bioenergy, 2010b) which is detrimental for saccharification when yeasts (such as *Saccharomyces cerevisiae* and *Candida intermedia*) are used for fermentation, as cooling to 25-38 °C (La Grange *et al.*, 2001; Leandro, Gonçalves, & Spencer-Martins, 2006; Toivari, Salusjärvi, Ruohonen, Salusja, & Penttila, 2004; Walfridsson *et al.*, 1996) after heating is costly (Abdel-Banat, Hoshida, Ano, Nonklang, & Akada, 2010). Therefore, both mesophilic and thermotolerant methods of culture were utilised.

Literature regarding isolation of microbes from biomass (e.g. *Miscanthus*), their identification and enzymatic characterisation in a biorefining context was scant. Most microorganisms that had been isolated related mostly to nitrogen fixing endophytes. Table 8.1 (Appendix 8.1.1) details reports relevant isolates to date, including characteristics and applications documented for each strain. The majority of bacteria isolated from *Miscanthus sp.* were shown to have biotechnological application, via lipase production (C. H. Liu, Lu, & Chang, 2006); soil remediation (Schallmey, Singh, & Ward, 2004); xylanase production (L. Liu *et al.*, 2013); cellulase production (Raghothama *et al.*, 2000); butanol synthesis (J. Lee *et al.*, 2012), and bioplastic degradation (Mergaert, Cnockaert, & Swings, 2002).

Many organisms, such as *Bacillus sp.* and *Clostridium thermocellum*, have genes encoding enzymes capable of breaking down complex plant cell wall carbohydrates. These long chain sugars, such as cellulose, hemicellulose along with starch, can be converted into monomers, central for metabolic function (Bertoldo & Antranikian, 2002; Gilbert & Hazlewood, 1993; Niehaus *et al.*, 1999). In turn this can result in the accumulation of metabolic products, such as ethanol, lactic acid and xylitol for biotechnological uses (Alves, Felipe, Silva, Silva, & Prata, 1998; Huang *et al.*, 2011; Lönn *et al.*, 2003; Su & Xu, 2014; Vintila *et al.*, 2010; Limin Wang *et al.*, 2010). Recent attention has turned towards thermophilic bacteria that produce such enzymes, and are able to withstand high temperature fermentation environments (Liang, Yesuf, Schmitt, Bender, & Bozzola, 2009; Madigan & Oren, 1999; Nakamura *et al.*, 1994; Niehaus *et al.*, 1999; Ozcan, Coskun, Ozcan, & Baylan, 2011; Rothschild & Mancinelli, 2001; Soemphol *et al.*, 2011). In this way the additional costs associated with maintaining temperature and pH could be reduced (Turner, Mamo, & Karlsson, 2007c). Therefore, it was hypothesised that isolates demonstrating activity across a carbon source, pH and temperature range could potentially be utilised in biorefining. This study focused on three carbon sources: cellulose (in the form of carboxymethylcellulose (CMC)), xylan (from Beechwood) and starch across a pH and temperature range. Though cellulose is the major sugar component of the plant cell wall (Table 1.1), hemicellulose depolymerisation is required for the full potential of the biomass to be reached (H. Yang, Yan, Chen, Lee, & Zheng, 2007; Yoshida *et al.*, 2008; Xiao Zhang, Tu, & Paice, 2011). Being insoluble at low temperatures in water, water solubility of hemicelluloses increases in response to increased temperature (Harmsen *et al.*, 2010). As such, bacterial isolates that can withstand a higher temperature, may

have a selective advantage with access to this carbohydrate source, and harbour the required hemicellulose depolymerising enzymes. The main constituent of hemicellulose in *Miscanthus* is the xylan molecule, which is predominantly composed of xylose (de Frias & Feng, 2013). Thus, by identifying bacterial isolates capable of cellulose, xylan and starch degradation novel enzymes and activities may be found for industrial saccharification.

In many bacterial systems secreted enzymes are used to degrade the backbone of xylan into short chain xylooligosaccharides (endo-acting) before it is transported into the cell, whereby it can be converted to xylose and metabolised (Damiano, Ward, Gomes, Alves-Prado, & Da Silva, 2006; Y. Han *et al.*, 2012; Honda & Kitaoka, 2004; Shin, McClendon, Vo, & Chen, 2010). Other microbes instead secrete a number of cellulases and hemicellulases (both endo- and exo-acting) to completely degrade the substrate before utilisation (Juturu & Wu, 2014; Nelson & Cox, 2005, p. 250; Pérez-Avalos *et al.*, 2008; Sánchez-Herrera *et al.*, 2007). Some systems use extracellular cell bound enzymes (Juturu & Wu, 2014; Kohring, Wiegel, & Mayer, 1990) or multi-enzyme complexes (cellulosomes) to degrade the crystalline polymer (Bayer, Shimon, Shoham, & Lamed, 1998; Fontes & Gilbert, 2010; Shulami *et al.*, 2007). Degradation of xylan, cellulose and starch has been demonstrated using basal medium agar supplemented with the target carbon source. Clear zones caused by enzymatic depolymerisation of the substrate have been used to assess activity (Bhadra, Rao, Singh, Sarkar, & Shivaji, 2008; Pointing, 1999). Carbon utilisation profiles were therefore based upon such methods. The size of halo in comparison to the amount of growth could also indicate if the enzymes were cytoplasmic, cell bound or secreted into the media (Bragger, Daniel,

Coolbear, & Morgan, 1989). Such secreted enzymes were also investigated using SDS-PAGE protein profiling (section 4.2.8).

There are multiple applications for xylanases, for example; food and beverage industries, feedstock improvement, lignocellulosic residue quality improvement (Pal & Khanum, 2010) and biobleaching (La Grange *et al.*, 2001). A cascade of enzymes is required to depolymerise xylan before it can be metabolised, with a by-product of ethanol (Javier *et al.*, 2013; La Grange *et al.*, 2001; Toivari *et al.*, 2004). However, the main focus of this study was monomer releasing enzymes, as D-glucose and D-xylose are major fermentable sugars, from a range of biomass via saccharolytic hydrolysis (Sedlak & Ho, 2004). Other molecules of interest can be found in plant biomass that can be fermented to produce useful chemicals, such as the degradation of starch to glucose which can be fermented to ethanol (Borglum, 1980; Chethana *et al.*, 2011). A range of industries use microbial starch depolymerising enzymes, such as brewing, baking, distilling (Chakraborty *et al.*, 2011), detergent production, pharmaceuticals, food processing (Chimata *et al.*, 2011), textile and paper manufacturing (R. Gupta, Gigras, Mohapatra, Goswami, & Chauhan, 2003; Hmidet *et al.*, 2008). Therefore, starch was included in the carbon utilisation analysis, though senesced *Miscanthus* has a low starch content (Lygin *et al.*, 2011; Purdy *et al.*, 2013). However, all three classes of enzyme would have industrial application by releasing monomeric sugars from pre-treated biomass, which could then be metabolised into products.

In order to isolate bacteria and enzymes that would be of use for saccharification or fermentation, screens were used to identify isolates with activity at: high pH, low pH, across a wide pH range, at high temperature, and across a carbon range with emphasis

on xylan. Isolates present on *Miscanthus* chip, which had been degrading in the field for 2.5 years, or soil surrounding *Miscanthus* plots were targeted. Polymerase Chain Reaction (PCR) analysis of isolates cultured from *Miscanthus* environments was performed for taxonomic identification, via 16S ribosomal RNA (rRNA) gene (Figure 2.1). Isolates were identified using 16S rRNA gene sequences, morphology and Gram staining. Morphological observations were performed in order to support the 16S identification and improve bacterial identification. Gram staining aided data from the literature and a complete bacterial assessment was gained and provided accurate identification.

The hypervariable regions found in small subunit rRNA that formed the 16S gene have been used extensively for the identification of bacteria (Ashelford, Chuzhanova, Fry, Jones, & Weightman, 2005; Bodilis, Nsigue-Meilo, Besaury, & Quillet, 2012; Chakravorty, Helb, Burday, Connell, & Alland, 2007; J. R. Hall *et al.*, 2008; Janda & Abbott, 2007; Suzuki & Giovannoni, 1996; The Walser Group, 2012). This technique was shown to be quicker and less cumbersome that a number of other methods that could be used, such as DNA-DNA hybridization, which is accurate especially for ambiguities (Janda & Abbott, 2007). Primers were designed to target either the whole 16S gene or a number of hypervariable regions within the gene (Figure 2.1) in order to identify bacterial species (J. C. Lee & Gutell, 2012; Noller, Stolk, Douthwaite, & Gutell, 1985). Eighteen different universal primer combinations were used to gain 16S rRNA sequences. Many combinations had been publicised, one of the most common was 8f and 1492R (Hanshew, Mason, Raffa, & Currie, 2013; Klindworth *et al.*, 2013; Sagaram

et al., 2009). Due to their prevalence in the literature, 8f and 1492R, and 63f and 1378R were the most utilised combinations.

This chapter reports a range of bacteria that are potentially able to break down feedstocks, such as *Miscanthus,* which could be utilised by industry in the future. A pH range was devised whereby activity could be assessed across the entire bacterial collection from as low as pH 4 to a high alkalinity. These were determined as: Starch – pH 4, 6.2, 8, 9; CMC – pH 4, 6.2, 8, 10; Xylan – pH 4, 6.2, 8.5 11. Mesophilic isolates also were assayed at pH 5.5 as an initial experiment, before the pH range was developed. Bacteria with activity at low and high pH, and high temperature, across a range of substrates were then selected for further analysis (Chapter 4).

3.2 Results

From *Miscanthus* chip and soil surrounding plots, 148 pure or mixed cultures were isolated, with multiple strains identified as the same species. Figure 3.1 described the range of bacteria isolated and the frequency each isolate appeared after 16S sequencing. Figure 3.1 shows that 42 different isolates were found, with 27 at species level, and 15 at a genus level. Appendix 1 (section 8.1.3) highlighted similarity scores from BLAST hits across two databases (NCBI BLAST and RDP) to gain final identifications (section 2.2.3.2). Scores should have had a cut off of 97% query coverage and maximum identity for a significant result. However results with lower homology were included as potential identification when nothing higher could be obtained.

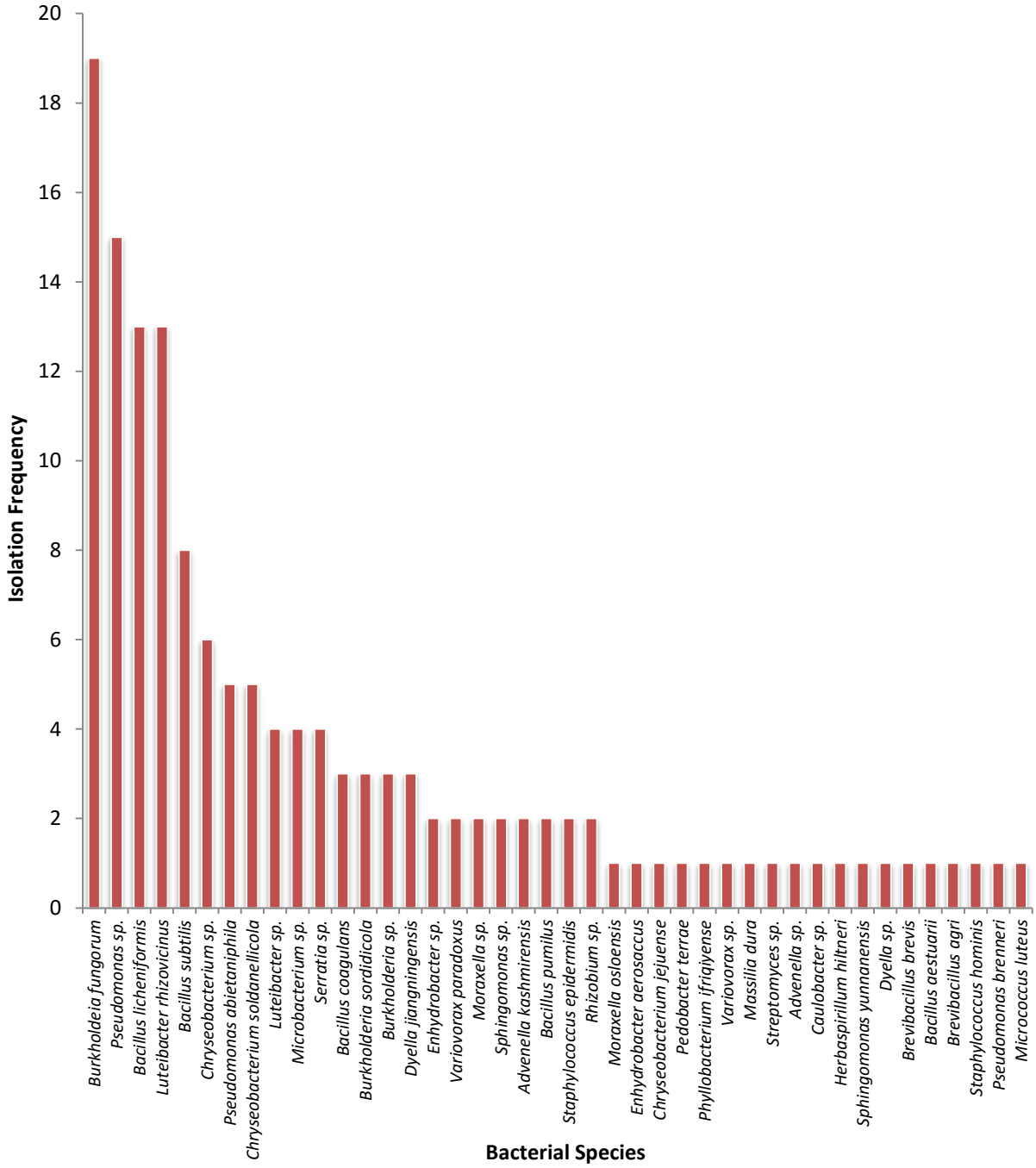


Figure 3.1 Range of bacteria isolated from *Miscanthus* chip and soil, and the frequency at which isolates appeared, based on best match homology to known 16S rRNA sequences.

Most common isolates were *Burkholderia fungorum* and *Pseudomonas sp.*. Of the 42 genus identified 19 isolates had low frequency whereas the remaining 129 had 2 or more isolates within the same species.

3.2.1 Bacterial identification by sequencing 16S rRNA PCR amplicons

Bacteria were identified using PCR whereby sections of the 16S rRNA gene were amplified and sequenced. A range of universal primer combinations were utilised (Table 2.1) to gain 16S rRNA identification of each colony. Preliminarily, 18 primer combinations were used to amplify the 16S gene from 4 soil bacterial isolates to find the most efficient pairs. The most efficient combinations were then utilised for the remaining bacterial collection. Amplification conditions were optimised throughout screens. Primer bias was almost removed as a number of different primer combinations were used when a sample could not be amplified.

From 148 isolates there were 42 genera and seven clades (Figure 3.2). *Bacillus sp.* all grouped into a clade bar *B. coagulans* ABR, which did not appear as closely related to the other *Bacillus sp.* whereby growth curves were investigated (Figure 3.6). *Pseudomonas sp.* ACH and ACO did not group with the other *Psuedomonas sp*.. Gram positive bacteria appeared to group together except where *Microbacterium sp.* (AEP, ADS, ADQ and AEU) and *Micrococcus luteus* ADM grouped with gram negatives *Dyella sp.* AFG, *Enhydrobacter sp.*/*Moraxella sp.* (ACB, ABF2 and ADK), *Herbaspirillum hiltneri* AEQ, *Massila dura* ADW, *Phyllobacterium ifriqiyense* ADT and *Pedobacter terrae* ACL. Gram positives grouped together bar where *Staphylococcus sp.* (M1, M2 and AITb) and *Streptococcus sp.* AEA grouped with Gram negative *Variovorax paradoxus* (ABI and ADE).

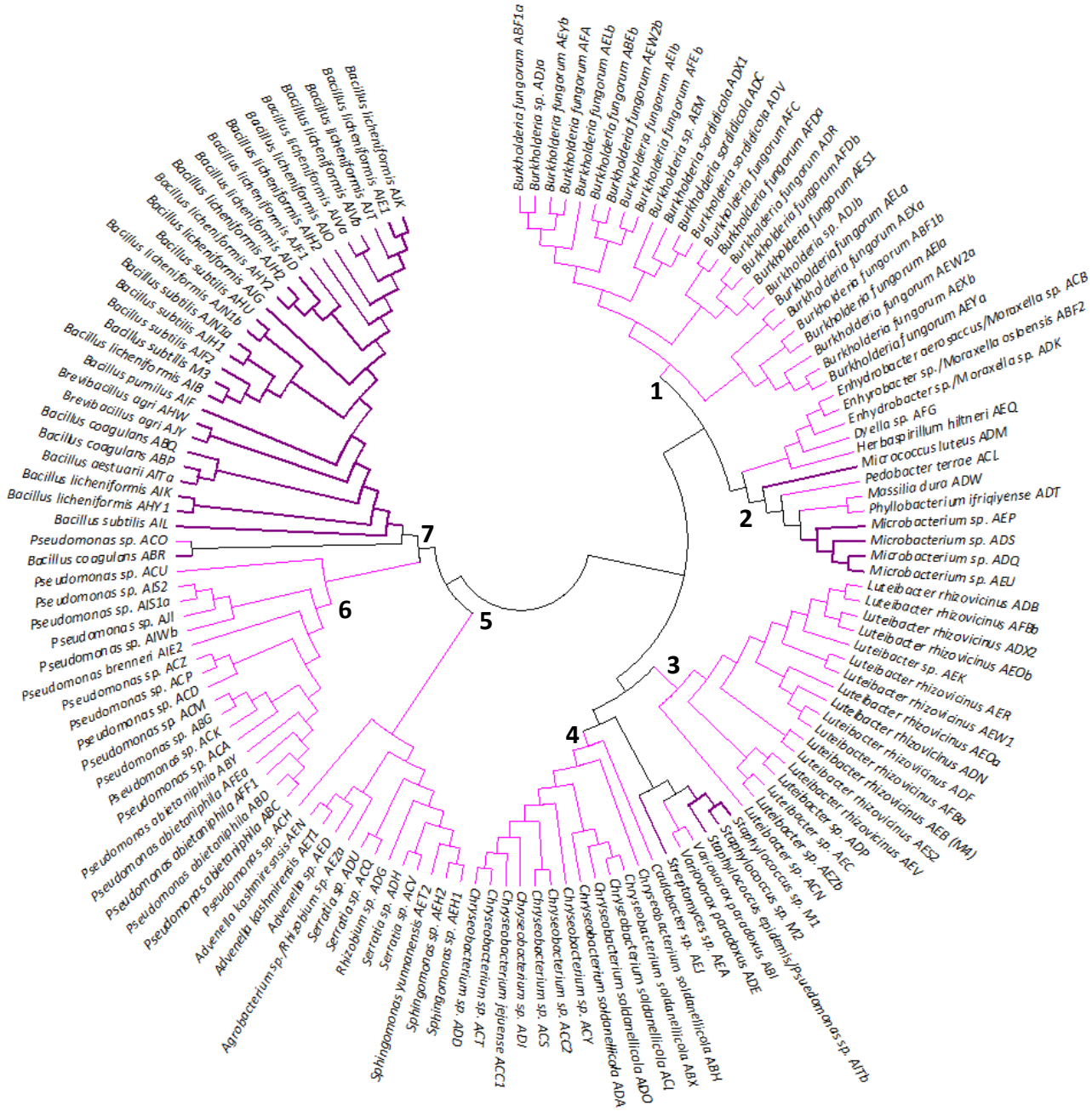

Burkholderia fungorum ABF1a
Burkholderia sp. ADJa
Burkholderia fungorum AEYb
Burkholderia fungorum AFA
Burkholderia fungorum AELb
Burkholderia fungorum ABEb
Burkholderia fungorum AEW2b
Burkholderia fungorum AFIb
Burkholderia fungorum AFEb
Burkholderia sp. AEM
Burkholderia sordidicola ADX1
Burkholderia sordidicola ADC
Burkholderia sordidicola ADV
Burkholderia fungorum AFC
Burkholderia fungorum AFDa
Burkholderia fungorum ADR
Burkholderia fungorum AFDb
Burkholderia fungorum AES1
Burkholderia sp. ADJb
Burkholderia fungorum AELa
Burkholderia fungorum AEXa
Burkholderia fungorum ABF1b
Burkholderia fungorum AEIa
Burkholderia fungorum AEW2a
Burkholderia fungorum AEXb
Burkholderia fungorum AEYa
Enhydrobacter aerosaccus/Moraxella sp. ACB
Enhydrobacter sp./Moraxella oslensis ABF2
Enhydrobacter sp./Moraxella sp. ADK
Dyella sp. AFG
Herbaspirillum hiltneri AEQ
Micrococcus luteus ADM
Pedobacter terrae ACL
Massilia dura ADW
Phyllobacterium ifriqiyense ADT
Microbacterium sp. AEP
Microbacterium sp. ADS
Microbacterium sp. ADQ
Microbacterium sp. AEU
Luteibacter rhizovicinus ADB
Luteibacter rhizovicinus AFBb
Luteibacter rhizovicinus ADX2
Luteibacter rhizovicinus AEOb
Luteibacter sp. AEK
Luteibacter rhizovicinus AER
Luteibacter rhizovicinus AEW1
Luteibacter rhizovicinus AEOa
Luteibacter rhizovicinus ADN
Luteibacter rhizovicinus ADF
Luteibacter rhizovicinus AEBa
Luteibacter rhizovicinus AES2
Luteibacter rhizovicinus AEB (M4)
Luteibacter rhizovicinus AEV
Luteibacter sp. ADP
Luteibacter sp. AEC
Luteibacter sp. AEZb
Luteibacter sp. ACN
Luteibacter sp. M1
Luteibacter sp. M2
Staphylococcus epidermis/Pseudomonas sp. AITb
Staphylococcus sp. ABI
Staphylococcus paradoxus ADE
Variovorax paradoxus ABJ
Streptomyces sp. AEJ
Caulobacter sp. AEA
Chryseobacterium soldanellicola ABH
Chryseobacterium soldanellicola ABX
Chryseobacterium soldanellicola ACL
Chryseobacterium soldanellicola ADO
Chryseobacterium soldanellicola ADA
Chryseobacterium sp. ACV
Chryseobacterium sp. ACC2
Chryseobacterium sp. ACS
Chryseobacterium sp. ADJ
Chryseobacterium jejuense ACC1
Chryseobacterium sp. ACT
Sphingomonas sp. AEH1
Sphingomonas sp. AEH2
Sphingomonas yunnanensis AET2
Serratia sp. ACV
Serratia sp. ADH
Serratia sp. ACQ
Rhizobium sp. ADU
Serratia sp. AEZb
Advenella sp. AED
Agrobacterium sp./Rhizobium sp. AETL
Advenella kashmirensis AEN
Advenella kashmirensis ACH
Pseudomonas sp. ABD
Pseudomonas abietaniphila ABC
Pseudomonas abietaniphila AFF1
Pseudomonas abietaniphila AFEa
Pseudomonas abietaniphila ABY
Pseudomonas sp. ACA
Pseudomonas sp. ACK
Pseudomonas sp. ABG
Pseudomonas sp. ACM
Pseudomonas sp. ACD
Pseudomonas sp. ACP
Pseudomonas sp. ACZ
Pseudomonas brenneri AIE2
Pseudomonas sp. AIWb
Pseudomonas sp. AJI
Pseudomonas sp. AIS1a
Pseudomonas sp. AIS2
Pseudomonas sp. ACU
Bacillus coagulans ABR
Pseudomonas sp. ACO
Bacillus subtilis AIL
Bacillus licheniformis AHY1
Bacillus licheniformis AIK
Bacillus aestuarii AIT a
Bacillus coagulans ABP
Bacillus coagulans ABQ
Bacillus coagulans AJY
Brevibacillus agri AHW
Brevibacillus agri AIF
Bacillus pumilus AIB
Bacillus licheniformis M3
Bacillus subtilis AIF2
Bacillus subtilis AIH1
Bacillus subtilis AJN1a
Bacillus subtilis AJN1b
Bacillus subtilis AHU
Bacillus licheniformis AJG
Bacillus licheniformis AHY2
Bacillus licheniformis AJD
Bacillus licheniformis AHF1
Bacillus licheniformis AHH2
Bacillus licheniformis AJO
Bacillus licheniformis AJV a
Bacillus licheniformis AJM b
Bacillus licheniformis AJE1
Bacillus licheniformis AJT
Bacillus licheniformis AJK
1
2
3
4
5
6
7

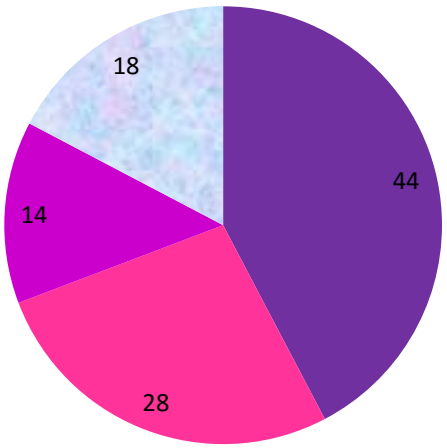
Figure 3.2 Neighbour-Joining tree of all 16S rRNA sequences from bacterial library.

The evolutionary history was inferred using the Neighbor-Joining method (Saitou & Nei, 1987). The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analysed (Felsenstein, 1985). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura, Nei, & Kumar, 2004) and are in the units of the number of base substitutions per site. The analysis involved 137 nucleotide sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 322 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013). 1 – *Burkholderia sp.* clade; 2 – mixed clade 1; 3 – *Luteibacter sp.* clade; 4 – *Chryseobacterium sp.* clade; 5 – mixed clade 2; 6 – *Pseudomonas sp.* clade, and 7 – *Bacillus sp*. clade. Bar isolates in less defined clades gram negative and gram positive strains grouped together. Isolates where 16S rRNA gene could not be amplified were omitted from this tree. Pink – Gram negative based on 16S rRNA gene identification; purple – Gram positive based on 16S rRNA gene identification. From 16S rRNA sequencing 38 sequences were homologous to Gram positive strains, whereas 110 sequences had homology to Gram negative strains.

3.2.2 Gram staining of mesophilic and thermotolerant isolates

The majority of isolates within the bacterial library were Gram positive (Figure 3.3, a) mesophiles: 42.3% and b) thermotolerant isolates: 73%). However there were 26.9% Gram negative isolates in the mesophilic collection, whereas 21.6% were Gram negative in the thermotolerant collection. There was only one isolate considered Gram variable, and one mixed colony within the thermotolerant collection. However within the mesophiles 17.3% were mixed colonies and 13.5% considered Gram variable. However, 16S rRNA identifications (Figure 3.2) showed that 74.3% of isolates were supposed to be Gram negative and 25.7% Gram positive.

a) Mesophiles



b) Thermotolerant isolates

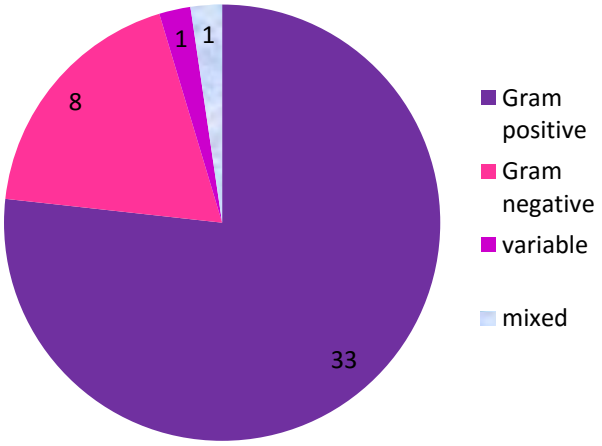


Figure 3.3 Proportion of Gram negative, Gram positive, Gram variable and mixed bacteria identified in the bacterial collection.

a) 44 isolates were seen to be Gram positive; 28 were observed to be Gram negative; 18 were seen to be mixed colonies, and 14 were thought to be Gram variable. b) 27 isolates were thought to be Gram positive; eight isolates were seen to be Gram negative; one isolate was considered Gram variable, and one isolate was observed to be a mixed colony.

Colonies grown on culture plates were used for Gram staining of mesophilic isolates. The 7 isolate Gram stains shown in

Figure 3.4 were found to have activity at low pH via carbon utilisation assays and were selected for whole genome sequencing (section 5.3). However isolates *a-d* were most likely mixed colonies (

Figure 3.4), though 16S primer bias meant that only one species per sample was identified due to oligonucleotides preferentially binding to specific 16S rRNA genes, not universally. Gram stain images were taken using a specialised camera attachment on the microscope which was linked to a laptop. Scales were set using ImageJ 1.46r (Wayne Rasband, National Institutes of Health) and the formula in Equation 3.1 to set parameters. The known distance was 4.1 μm in 100 pixels. Many of the bacteria are less than 1μm in length (

Figure 3.4).

Equation 3.1 Parameter settings for scaling of Gram stain photos

$$\frac{\text{length in pixels x 10000 um/cm}}{\text{resolution in pixels/cm x magnification}}$$

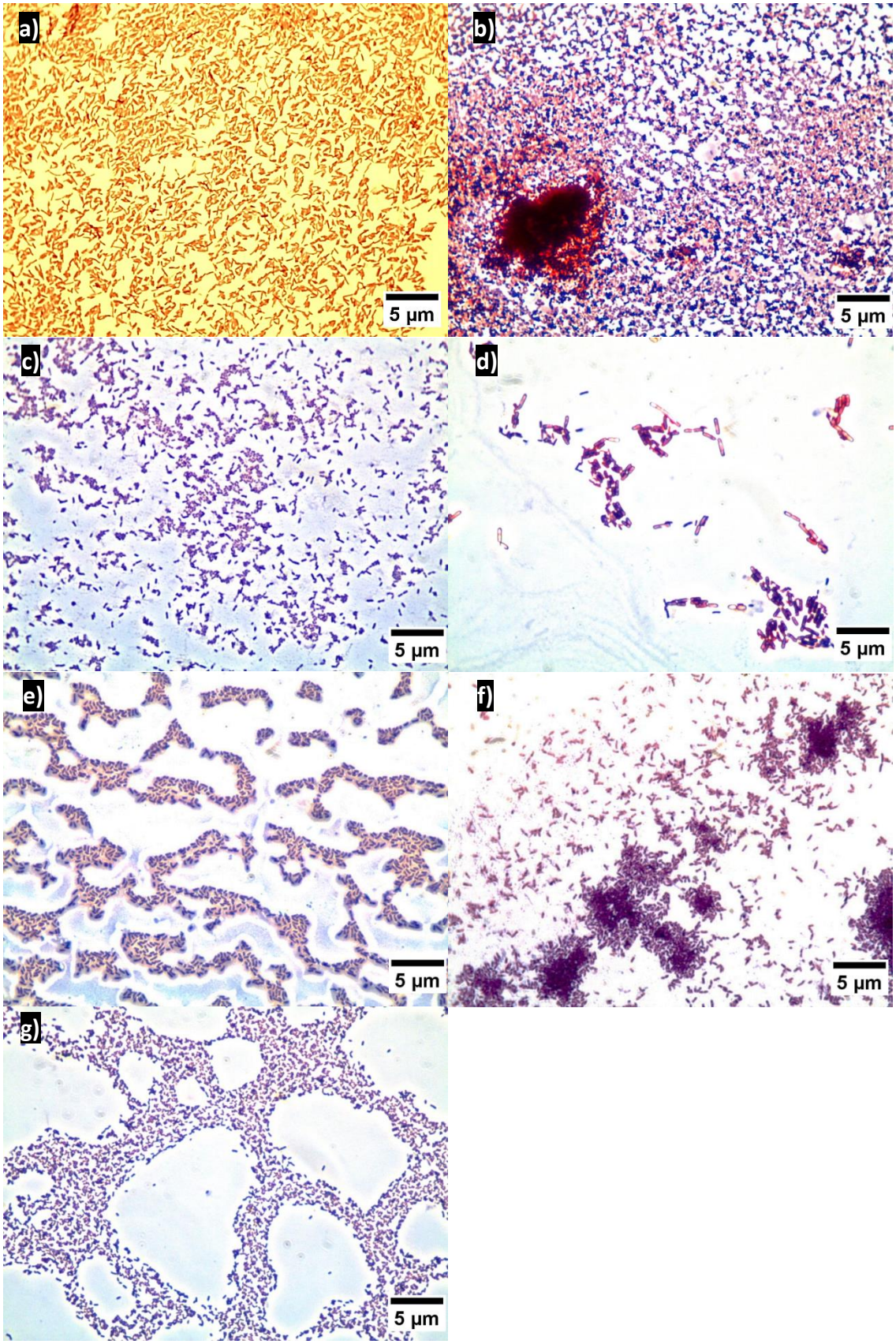

a)
5 µm
b)
5 µm
c)
5 µm
d)
5 µm
e)
5 µm
f)
5 µm
g)
5 µm

Figure 3.4 Gram staining photos at x100 oil immersion lens for seven mesophilic bacterial isolates.

Isolates were later taken forward for whole genome sequencing based on their carbon utilisation profiles. Gram stains were conducted from agar plate growth for mesophilic isolates. a) *Burkholderia fungorum* AFDa (mixed); b) *Luteibacter rhizovicinus* AEOb (mixed); c) *Burkholderia sp.* AEW2a (mixed); d) *Burkholderia fungorum* AEIb; e) *Sphingomonas sp.* AEH1; f) *Burkholderia fungorum* AELa; g) *Agrobacterium sp./Rhizobium sp.* AEZa (variable)

The use of liquid culture for thermophilic Gram stains (

Figure 3.5) meant that the bacterial cells were less crowded than in mesophilic isolates staining (

Figure 3.4). There was also a proportion of dye still present and there were shadows where cells had been fixed to the slide. Most isolates from this collection were rod shaped (

Figure 3.5); this was not unusual as many soil and plant associated bacteria are generally of this shape (W. De Boer, Folman, Summerbell, & Boddy, 2005; Glick, 2012; Kloepper, Ryu, & Zhang, 2004). Many of the bacteria present in the thermophilic collection had cells that were larger than 1 μm and there were fewer mixed samples. There was a range of Gram positive and negative bacteria in this collection and only three isolates were identified as coccidian. All isolates were able to grow on Nutrient agar though some were isolated from MRS agar. However not all isolates were able to grow on MRS agar, of which the components were more specific for lactic acid producing bacteria.

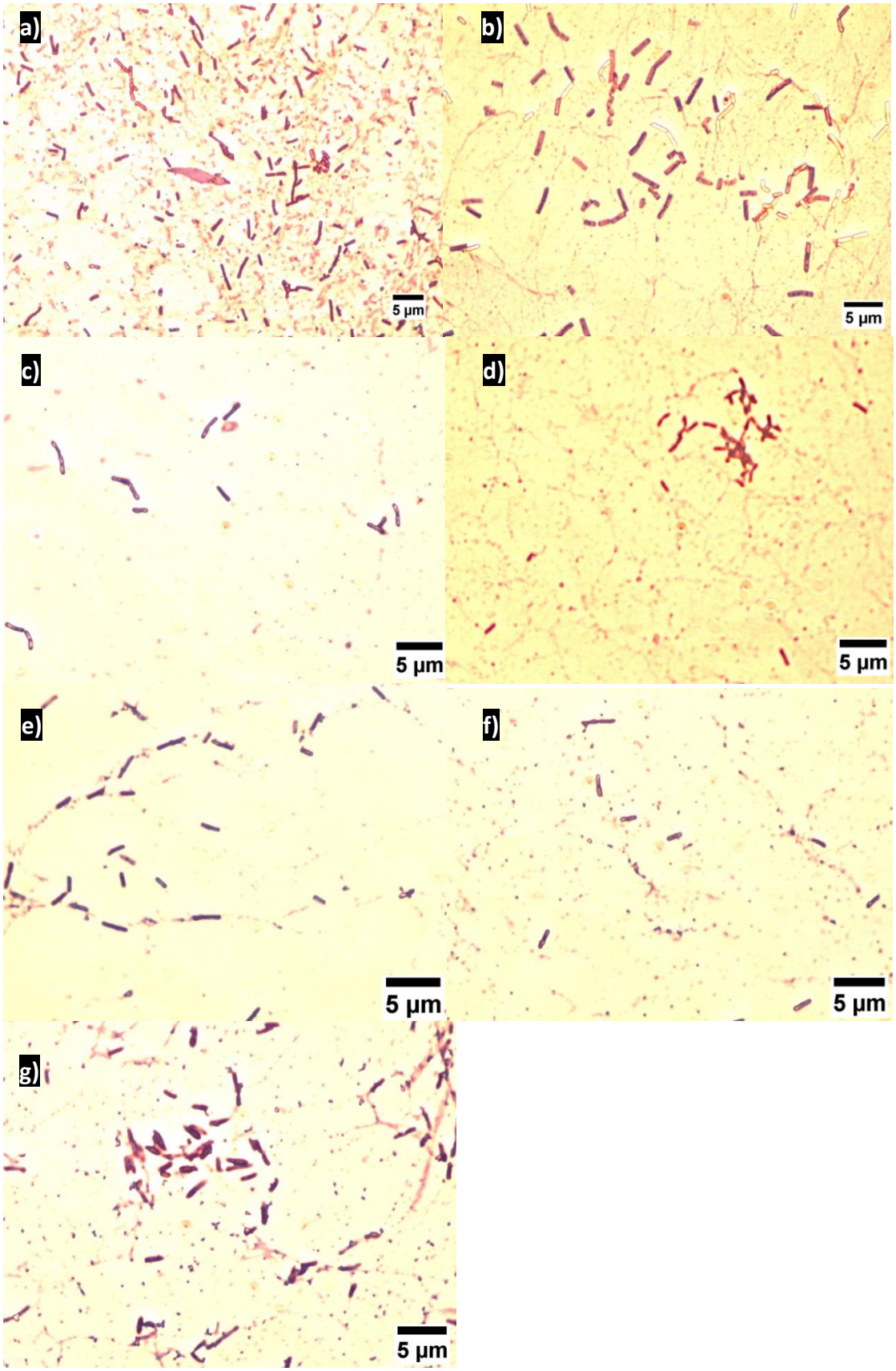

a)
5 µm
b)
5 µm
c)
5 µm
d)
5 µm
e)
5 µm
f)
5 µm
g)
5 µm

Figure 3.5 Gram stained images used for identification of bacterial isolates from liquid culture, within the thermophilic collection.

Stains were photographed via x100 oil immersion lens for seven bacterial isolates that were later taken forward for whole genome sequencing based on their performance on carbon utilisation assays. a) AJF2 – *B. subtilis*; b) AIB – *B. licheniformis* (variable); c) AJK – putative *B. licheniformis*; d) AJF1 – *B. licheniformis*; e) AIF – putative *B. stratosphericus*; f) AHY1 – putative *B. licheniformis*; g) AHW – *B. agri*.

3.2.3 Morphology of bacterial library

Isolates within the bacterial library were analysed morphologically using a dissection microscope and folding magnifier (10x) to view colonies on agar plates. Morphology was analysed based on colour, surface, form, elevation and margin of each colony. Based on 16S rRNA identification 10 isolates could not be determined (ND) and eight could not be cultured for morphological assessment so were also ND. When cultured, six different colours of colony were observed, with 55% cream, 19% yellow, 10.9% white, 6.8% orange, 1.4% grey white and only one isolate pink. Surface appearance was mainly shiny (87%) with only 7.5% having a dry exterior. Colonies had four forms: punctiform, circular, filamentous and irregular. Most common form was circular (51%), 38% were punctiform and 5.4% either irregular or filamentous. Colony elevation on agar plates were largely raised (37.4%) or flat (36.7%), only 17% were convex, 2.7% pulvinate and one isolate crateriform. 73% of isolates had entire margins; 12% were undulate; 6% had erose margins, and 3% either lobate or filamentous. Almost 58% of isolates in the bacterial library were opaque. The remainder were translucent though nearly 5% were opaque with translucent edges.

Bacteria that were isolated at room temperature (104 isolates) were grouped from clade 1-5 (Appendix 1: 8.1.2 and Figure 3.2). Of these, 42.3% of isolates were Gram positive, 26.9% Gram negative, 17.3% mixed colonies and 13.5% classed as Gram variable. Only 14.4% were coccidian, 8.7% mixed rods and cocci, with the remaining majority of isolates rod shaped (76.9%). The highest proportion of isolates were cream coloured (37.5%); 26.9% were yellow; 17.3% white; 9.6% orange; 1.9% grey white, and one isolate was pink. Due to lack of growth for morphological analysis 7.7% of colonies

were classed as not determined. The surface of most colonies (86.5%) was shiny, only 5.8% were dry. Form was split between punctiform (50%) and circular (42.3%). Elevation was observed to be mostly raised (40.4%); followed by 38.5% flat; 10.6% convex; 1.9% pulvinate, and only one colony crateriform. Margins of colonies were most frequently entire (81.7%); 6.7% erose; 2.9% undulate; 1.9% filamentous, with one lobate colony. Translucency of colonies was split between opaque (44.2%) and translucent (41.3%); were 6.7% opaque with a translucent edge.

Bacteria that were isolated at high temperatures of 40-70 °C (43 isolates), were mostly found in clade 6 and 7 (Figure 3.2 and 8.1.2). The majority of isolates were Gram positive (73%) with just 22% gram negative; one isolate was variable, and one was identified as a mixed colony. Of the thermophiles, 88% were rod shaped, 9% cocci and one isolate was a mixed colony of both rods and cocci. However, six isolates were not identifiable through 16S rRNA analysis. Though all isolates in this collection were cream in colour there were differences in other aspects of their morphology such as surface appearance where 11% were observed to be dry, where as 88% were shiny. Colony form was less diverse than mesophilic isolates; 72% were circular, 16% were irregular, 9% punctiform, and 2% filamentous. Elevation of colonies was more widespread with 32.6% flat, 32.6% convex, 30.2% raised and 4.7% pulvinate. The majority of isolates margins were entire (55.8%), though 34.9% were undulate, 4.7% erose, 2.3% erose and 2.3% lobate. Most isolates were opaque (90.7%), only 9.3% were transparent.

3.2.4 Growth curves analysis of *Bacillus coagulans* strains

B. coagulans strains ABP, ABQ and ABR were isolated from soil surrounding *Miscanthus* plots. Due to the odd grouping of ABR in Figure 3.2, whereby ABP and ABQ were more closely related, strains were subject to growth curve analysis. Strains were isolated at 40 °C on MRS agar at pH 5.5. After 16S rRNA identification and morphological analysis strains were subject to growth curve analysis to identify differences in growth pattern. Isolates were grown in MRS broth at pH 6.2 and 40 °C. The pH of the supernatant was also tested to identify metabolic by-product formation for carbon sources with MRS broth. Growth was assessed using the optical density (OD_{600}) of 1.5 mL of liquid culture every hour (Figure 3.6). The pH was taken at the end of the experiment at 12 hours growth and after 25 hours growth.

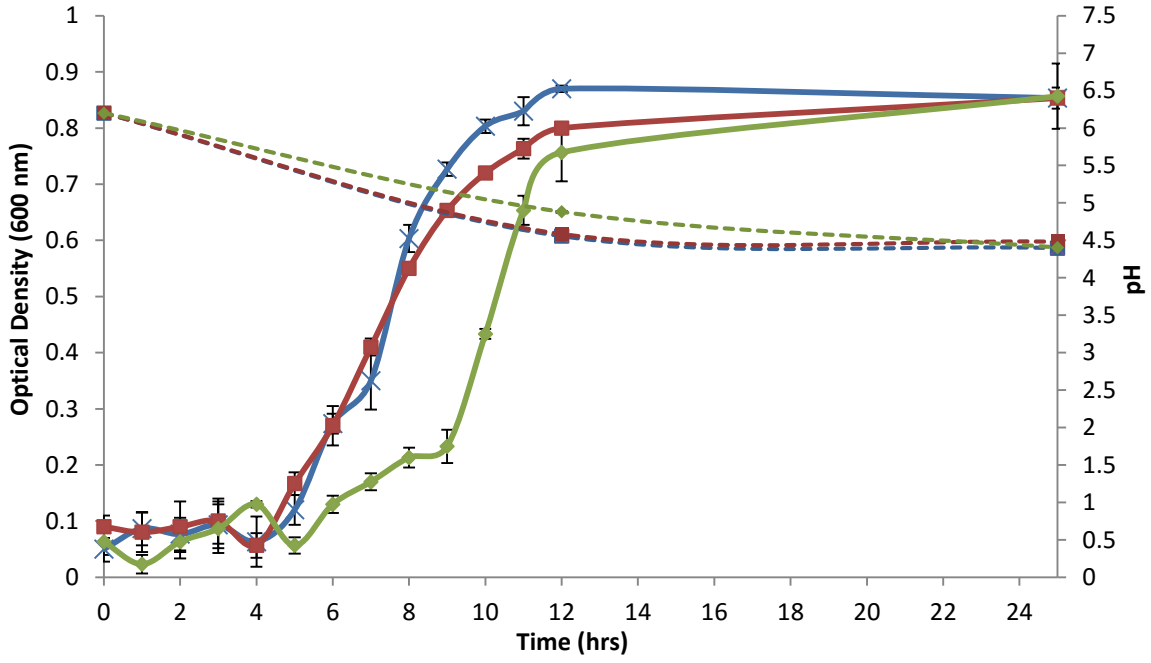


Figure 3.6 Growth curve of three *Bacillus coagulans* strains plus supernatant pH.

Readings over 12 hours plus a final reading after 25 hours growth to confirm plateau. *B. coagulans* ABP (x); *B. coagulans* ABQ (■), and *B. coagulans* ABR (♦); n=3. Dashed lines indicate changes in pH from initial readings to final readings at 12 hours and 24 hours. Standard deviation error bars are shown.

3.2.5 Carbon utilisation profiles

Both mesophilic and thermotolerant isolates from *Miscanthus* chip were assessed on their ability to deconstruct plant cell wall components (cellulose, xylan and starch) at a range of temperatures and pH. Simple plate based assays were used to provide quick visual results which would aid selection process for isolates of interest (Pointing, 1999). Isolates selected for next generation whole genome sequencing (Table 5.1) based on their efficacy at extremes of pH and high temperatures on xylan agar were presented here. As described in Materials and Methods (2.2.1), bacteria isolated from *Miscanthus* chip were inoculated onto three different carbon sources at different pH values (Appendix 8.1.3).

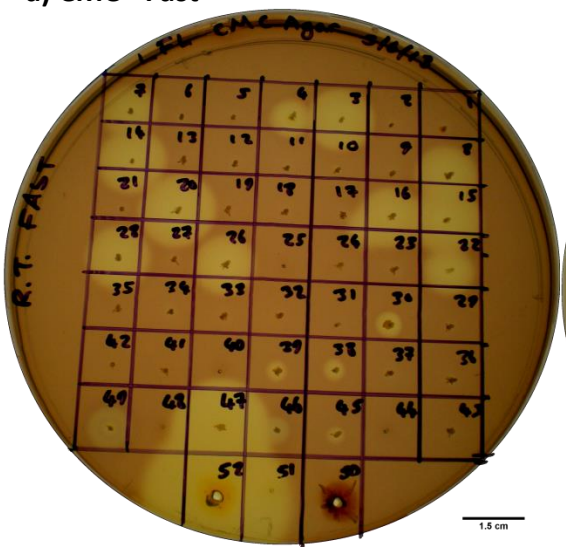
3.2.5.1 Mesophilic isolates

Mesophilic isolates were split into "fast growing" and "slow growing" to avoid colony overgrowth. Thermotolerant isolates were restricted to one plate, as they were less numerous and grew overnight. For most isolates across the three carbon sources optimum pH was 5.5-8, and activity was measured on a scale of 0-6. Approximate measurements (in mm) of halos were assigned to each number in the scale: 0 showed no activity halo; 1 represented 1-2 mm halo; 2 was 3-5 mm halo; 3 was 5-7 mm; 4 represented 8-10 mm halo; 5 was 11-13 mm halo, and 6 showed activity of 14 mm or more (where the halo expanded out of the grid square). Figure 3.7 shows the carbon utilisation assays results where each grid square contained a single isolate inoculated into the media using a sterile pipette tip (Table 3.1). Very clear halos could be seen on CMC and starch plates, whereas halos on xylan agar were less defined. The controls for each plate were overloaded but did not interfere with the assays.

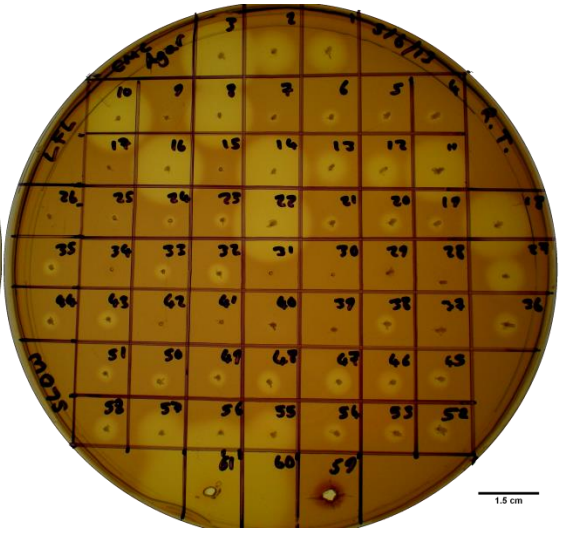
Mesophilic plates were incubated in a plastic box (at room temperature) for two nights for sufficient colony growth and detection of enzymatic activity. Activity halos are shown across the three carbon source assays in Figure 3.7. Starch and CMC degradation assays had very clear halos. However, the xylan plates had diffused halos, which could be due to the method of exo-enzymatic action (1.3.3 and 3.1). Each plate was inoculated with two negative controls; 2 μL of deionised water and nutrient broth respectively, which were inoculated in the same way as cultures. Clear halos were not observed from negative controls.

The relative activity of mesophilic isolates on xylan agar across a range of pHs was the main focus of the screens. The most interesting of these isolates were AEH1, AEH2, AEIb, AELa, AEOb, AEW2a, AEZa and AFDa as they demonstrated activity at pH 4 (Figure 3.8). Six of these isolates (bar AEH2 and AEIb) were taken forward for whole genome sequencing (section 5.2) due to their efficacy for low pH. These six isolates had activity up to pH 8.5, thereby demonstrating activity across a broad pH range. However, AEH2 did not grow when cultured from glycerol and therefore could not be used. None of the room temperature isolates were able to break down xylan at pH 11; few such as ACC1 and ACC2 had very high activity at pH 8.5. However, four isolates were not able to degrade xylan at any pH tested, and 23 were only able to break down xylan at a single pH.

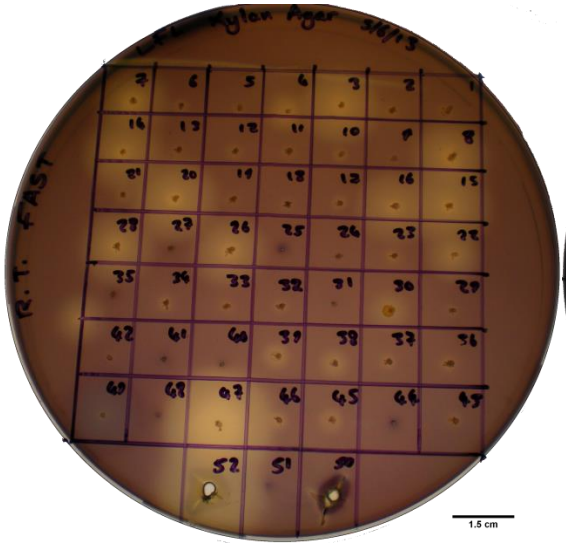
a) CMC - Fast



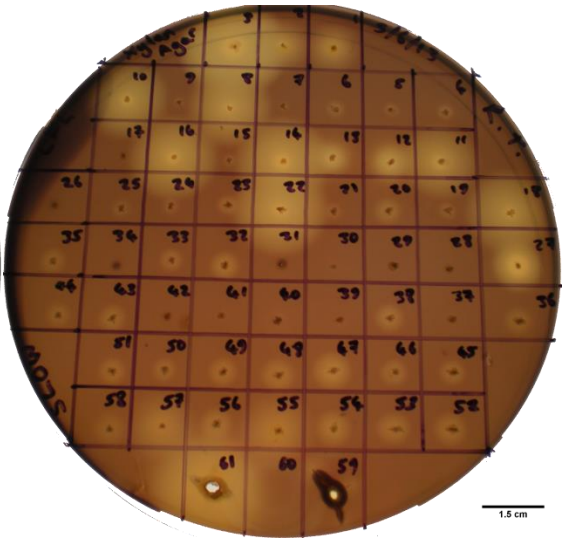
b) CMC - Slow



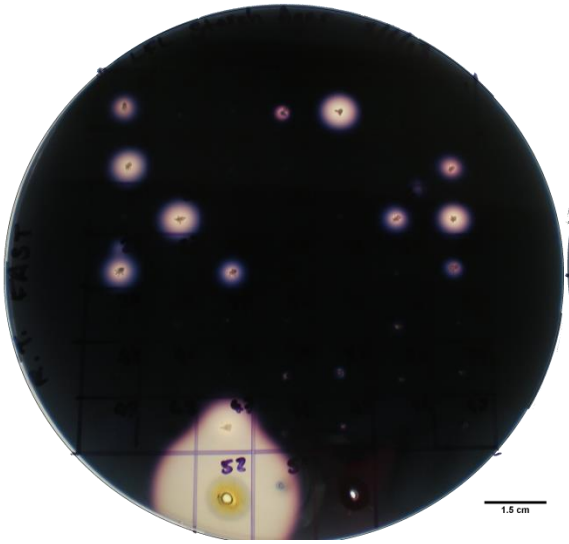
c) Xylan - Fast



d) Xylan - Slow



e) Starch - Fast



f) Starch - Slow

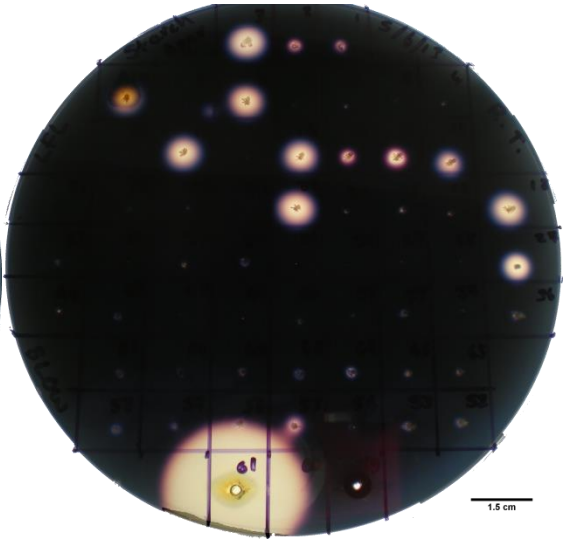


Figure 3.7 Mesophilic carbon utilisation plate assays of bacterial isolates split into fast and slow growers (identification table shown in **Table 3.1**).

Each square of the grid had a different isolate inoculated using a pipette tip. All plates were at pH 6.2 and incubated at room temperature for two days. Clear halo formation after iodine staining was used to assess enzymatic activity. *a) Carboxymethylcellulose (CMC) agar fast growing plate; b) CMC agar slow growing plate; c) xylan agar fast growing plate; d) xylan agar slow growing plate; e) starch agar fast growing plate; f) starch agar slow growing plate*. Positive controls: NS22086 Cellulase Complex; activity – 1000 BHU(2)/g, density – 1.15g/mL, pH optimum – 5.0-5.5, temperature optimum – 45-50 °C, dosage – 1-5% w/w total solids. NS22083 Xylanase; activity – 2500 fungal xylanase unit/g, density – 1.09 g/mL, pH optimum – 4.5-6.0, temperature optimum – 35-55 °C, dosage – 0.05-0.25% w/w total solids. Fisher bacterial alpha-amylase; 50-100 units/mg, decomposition temperature – 85 °C.

Table 3.1 Identification table for mesophilic carbon utilisation assays shown in Figure 3.7.

Isolates are recorded by their three letter identification code. a) Fast growing isolates; b) slow growing isolates.

a)

1 ABG	2 ABI	3 ABX	4 ABY	5 ACA	6 ACD	7 ACH
8 ACI	9 ACK	10 ACM	11 ACO	12 ACP	13 ACQ	14 ACS
15 ACT	16 ACU	17 ACV	18 ACW	19 ACX	20 ACY	21 ACZ
22 ADA	23 ADE	24 ADF	25 ADH	26 ADM	27 ADN	28 ADP
29 ADQ	30 ADS	31 ADT	32 ADU	33 ADW	34 AED	35 AEK
36 AEM	37 AEN	38 AEOa	39 AEOb	40 AEP	41 AER	42 AEU
43 AEV	44 AFDa	45 AFDb	46 AFG	47 M1	48 M2	49 M3
		50 dH_2O	51 -ve	52 +ve		

b)

		1 ABC	2 ABD	3 ABEa			
4 ABEb	5 ABF	6 ABF	7 ABF2	8 ABH	9 ACB	10 ACC1	
11 ACC2	12 ACL	13 ACN	14 ADB	15 ADC	16 ADD	17 ADG	
18 ADI	19 ADJa	20 ADJb	21 ADK	22 ADO	23 ADR	24 ADV	25 ADX1	26 ADX2
27 AEA	28 M4	29 AEC	30 AEH1	31 AEH2	32 AEIa	33 AEIb	34 AEJ	35 AELa
36 AELb	37 AEQ	38 AES1	39 AES2	40 AET1	41 AET2	42 AEW1	43 AEW2a	44 AEW2b
45 AEXa	46 AEXb	47 AEYa	48 AEYb	49 AEZa	50 AEZb	51 AFA		
52 AFBa	53 AFBb	54 AFC	55 AFEa	56 AFEb	57 AFF1	58 AFF2		
		59 dH_2O	60 -ve	61 +ve				

Relative activity of each isolate on CMC agar (n=2) was highest at pH 5.5 and 6.2, indicating that these isolates could tolerate slightly acidic environments. Many isolates also had activity at pH 8, but no isolates had activity at the extremes of the assay (pH 10 and pH 4) and were therefore not able to withstand more extreme environments. Relatively high activity on cellulose agar, all occurring on pH 5.5 or 6.2 assays was observed for 17 isolates (ABEa, ABH, ABX, ACC1, ACC2, ACI, ACS, ACT, ACU, ACY, ADA, ADB, ADD, ADI, ADM, AFEa and M1). ABF1a, ABF1b, AED and AFDb had higher relative activity at pH 8 than the other pHs tested. However, as none of the isolates had activity at pH 4 or 10 and were only tested at room temperature they were not taken into consideration for further testing. Activity at one pH (generally 5.5) occurred for 33 isolates and 10 had no activity at all across the pH range.

A larger proportion of isolates were able to break down starch at the highest and lowest pH. In total, 26 isolates were able utilise starch at pH 4 and 15 at pH 9, with a single isolate (AEOb) with activity across the whole pH range. Isolates AELa and AEYa have the best activity on pH 4 starch agar. AEOb and AELa were chosen for further xylanase characterisation (section 5.4.3). There were 27 isolates only able to utilise starch at a single pH (generally pH 5.5 or 6.2), and six had no activity at all across the pH range. *B. coagulans* ABQ was chosen for further enzymatic characterisation (section 4.2) due to its ability to depolymerise the starch polymer.

Though up to 14 isolates had high activity on CMC, xylan and starch, isolates chosen for whole genome sequencing appeared to have low activity. However, the 6 selected isolates (AEH1, AELa, AEOb, AEW2a, AEZa and AFDa) were chosen due to the efficacy for low pH on xylan and starch agar at mesophilic temperatures (Figure 3.8). Such

enzymatic activity could aid saccharification of biomass before fermentation via yeasts (section 1.3).

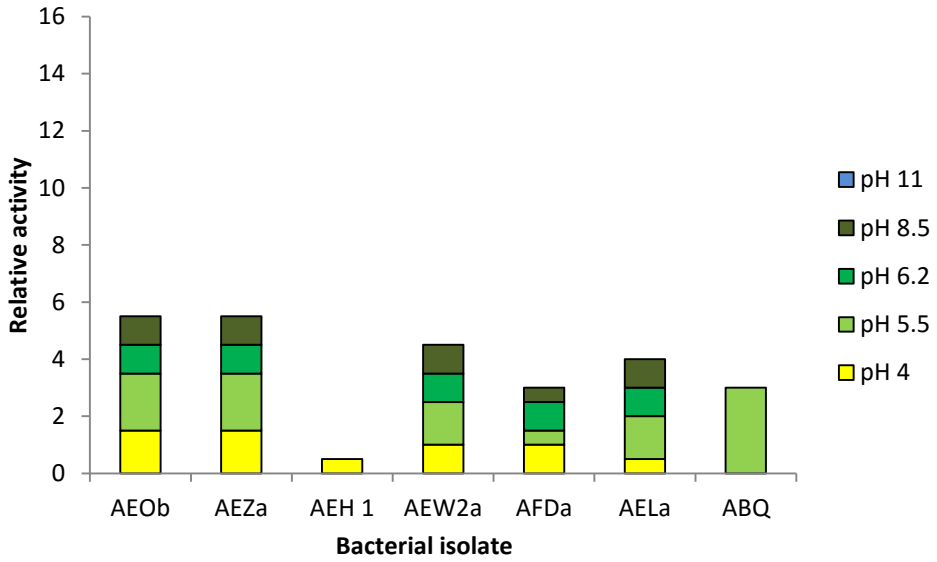


Figure 3.8 Carbon utilisation profiles of mesophilic bacterial isolates that were active at low pH (n=2).

AEOb – potential *L. rhizovicinus*; AEZa – potential *Luteibacter sp.;* AEH1 – potential *Sphingomonas sp*.; AEW2a – potential *B. fungorum*; AFDa – potential *B. fungorum*; AELa – potential *B. fungorum*, and ABQ – potential *B. coagulans*.

3.2.5.2 Thermotolerant isolates

Carbon utilisation assays were also used to assess thermotolerant bacterial enzyme activity on cellulose, xylan and starch (as the sole carbon source). Isolates were assayed across a temperature and pH range in conjunction with each carbon source. The xylan utilisation assays are shown in Figure 3.9 and Figure 3.10, and isolates are listed in Table 3.2 to identify the location of each bacterial colony. To assess isolate performance, and to narrow down selection for protein characterisation and next generation sequencing, four pH values were tested at four temperatures. Figure 3.9 demonstrates the temperature range at a single pH (pH 8.5); here activity was highest at 50 °C but declined as the temperature increased. However, a number of isolates

retained activity at 70 °C, though to a lower degree. It is also interesting to note the low activity of the commercial mixture that was used as a positive control (grid number 40), as the optimal pH was pH 4.5-6.0 though optimum temperature was 35-55 °C (Novozymes Bioenergy, 2010b). Activity and growth were measured in a scale of 0-6 (3.2.5). Plates were incubated at each temperature overnight, washed and stained with iodine the following day. Isolates AIB, AJT, AJN1b, AIL, AJH1, AHU, AJN1a, AJF2, AJG and AJK were chosen for further endoglucanase analysis (4.2.8) and whole genome analysis (5.2) as they had activity at pH 10 on CMC agar and pH 11 on xylan.

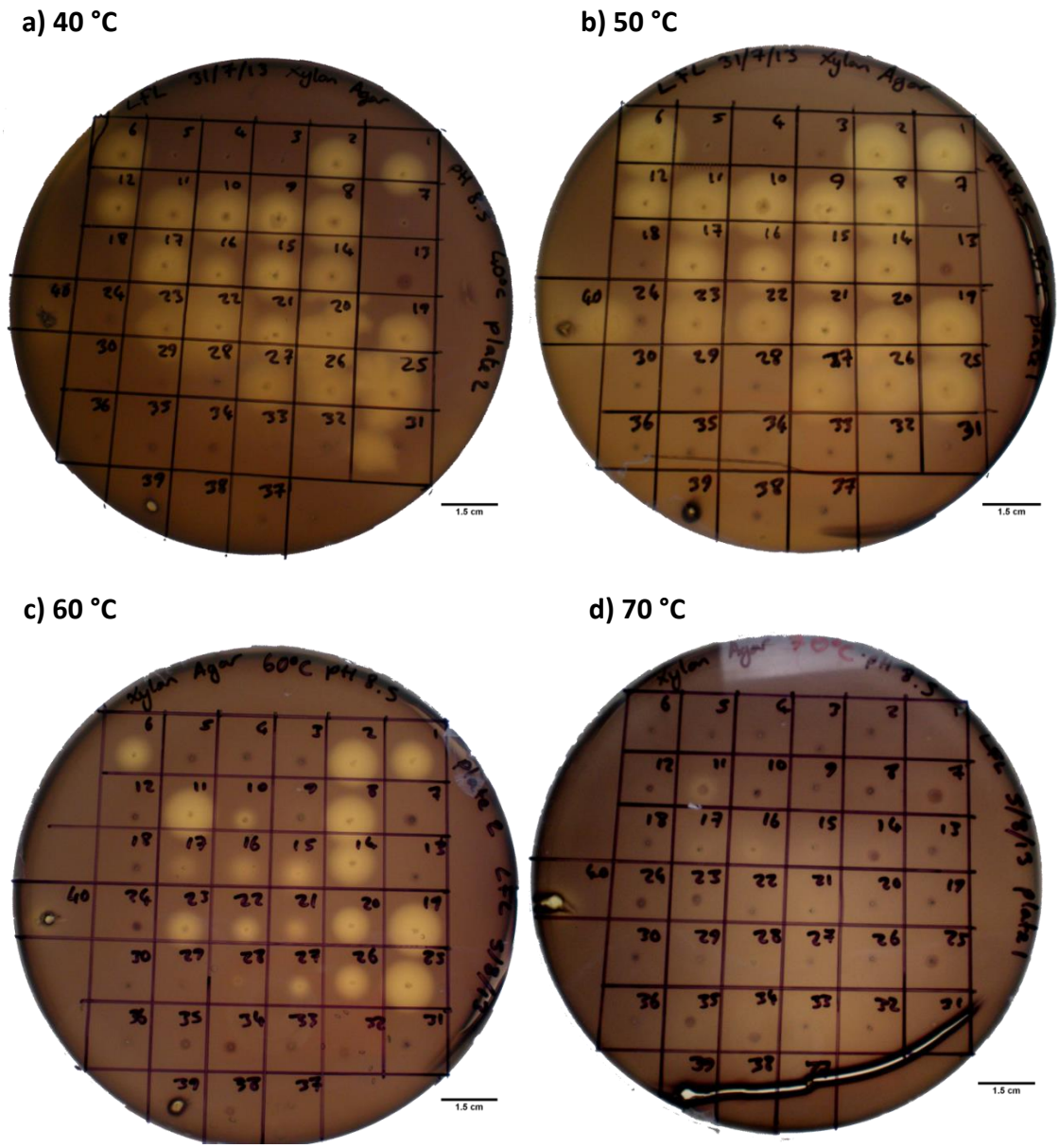


Figure 3.9 Thermotolerant bacterial isolates activity on xylan utilisation assay at pH 8.5 across temperature range 40-70 °C.

Positive controls: NS22086 Cellulase Complex; activity – 1000 BHU(2)/g, density – 1.15g/mL, pH optimum – 5.0-5.5, temperature optimum – 45-50 °C, dosage – 1-5% w/w total solids. NS22083 Xylanase; activity – 2500 fungal xylanase unit/g, density – 1.09 g/mL, pH optimum – 4.5-6.0, temperature optimum – 35-55 °C, dosage – 0.05-0.25% w/w total solids. Fisher bacterial alpha-amylase; 50-100 units/mg, decomposition temperature – 85 °C.

Figure 3.10 shows the pH range tested at a fixed temperature of 40 °C. Here, pH 6.2 and 8.5 were seen to have had the most activity. No activity was observed at pH 4 for any of the isolates inoculated, whereas the Novozymes control had a halo of xylan depolymerisation. In plate "b)" there was over growth of isolate 20 (AHY2 – potential *Bacillus licheniformis*), whereby the surrounding isolates results were less clear. However, both isolate 26 (AJVa - also potential *Bacillus licheniformis*) and 27 (AJVb – also potential *Bacillus licheniformis*) had an inhibitory effect of the outgrowth of isolate 20. The media in plate "a)" was softer than the other plates due to the low pH (AgarGel, 2003), therefore when the plate was blotted parts of the plate surface were removed leaving the pattern shown in Figure 3.10 a). Plates across the range of carbon sources showed similar results, whereby the extremes of pH (lowest: 4 – all carbon sources and highest: 9 – starch, 10 – CMC, 11 - xylan) had little or no activity. The either cells could not survive the conditions or the enzymes were denatured. However, the bacterial growth also decreased as the pH and temperature increased.

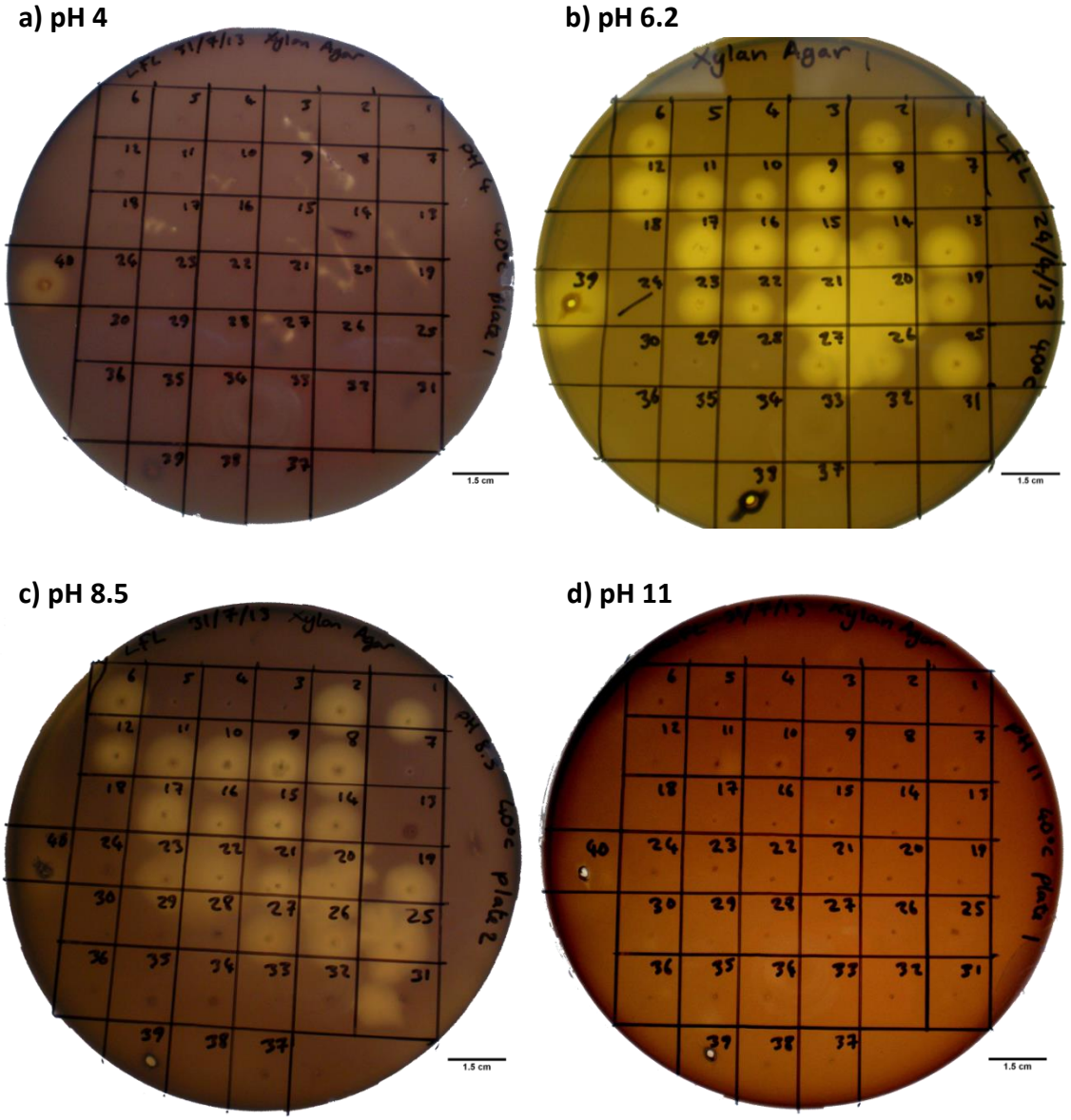


Figure 3.10 Thermotolerant bacterial isolates activity on xylan utilisation assay across pH range (pH 4-11) at 40 °C.

Positive controls: NS22086 Cellulase Complex; activity – 1000 BHU(2)/g, density – 1.15g/mL, pH optimum – 5.0-5.5, temperature optimum – 45-50 °C, dosage – 1-5% w/w total solids. NS22083 Xylanase; activity – 2500 fungal xylanase unit/g, density – 1.09 g/mL, pH optimum – 4.5-6.0, temperature optimum – 35-55 °C, dosage – 0.05-0.25% w/w total solids. Fisher bacterial alpha-amylase; 50-100 units/mg, decomposition temperature – 85 °C.

Table 3.2 Identification table for thermophilic carbon utilisation assays.

Plate b) from Figure 3.10 had AIWa in grid number 36, AIWb in 37 and the negative broth control in 38, however AIWa could not be cultured for the remainder of the plates. *Grid numbers 1-14 were isolated at 40 °C; 15-27 at 50 °C; 28-35 at 60 °C, and 36 at 70 °C. NB –ve - Nutrient broth negative control; MRS –ve – MRS broth negative control; dH2O –ve – distilled water negative control; +ve – xylanase (Novozymes, NS22083), cellulase Novozymes, NS22086) or alpha amylase (Fisher, A/7190/44) positive control.*

1 AID	2 AIE1	3 AIE2	4 AIF	5 AIH1	6 AIH2	
7 AIK	8 AJF1	9 AJF2	10 AJG	11 AJH1	12 AJH2	
13 AJI	14 AJK	15 AJN1a	16 AJN1b	17 AHU	18 AHW	
19 AHY1	20 AHY2	21 AIB	22 AIL	23 AJO	24 AJP	40 +ve
25 AJT	26 AJVa	27 AJVb	28 AIS1a	29 AIS1b	30 AIS2	
31 AITa	32 AITb	33 AIU	34 AJY	35 AKE	36 AIWb	
		37 NB -ve	38 MRS -ve	39 dH_2O -ve		

Enzymatic activities on xylan agar of each isolate chosen for whole genome sequencing are presented here. For further details of these isolates activity across CMC and starch, along with all other isolates tested see Appendix 8.1.3, Table 8.4 and Table 8.5. Isolate AHU (Figure 3.11) was identified by 16S rRNA sequencing to be *Bacillus subtilis*, concurring with the Gram stain as straight rods (Holt, Krieg, Sneath, Staley, & Williams, 2000). AHU was originally isolated at 50 °C though this isolate more active on CMC and starch at 40 °C but slightly less so on xylan. It could also be seen that this isolate was able to survive at high temperatures (up to 70 °C). Though AHU was not able to grow at pH 4 on any carbon source tested it had activity at pH 11 from 40-60 °C.

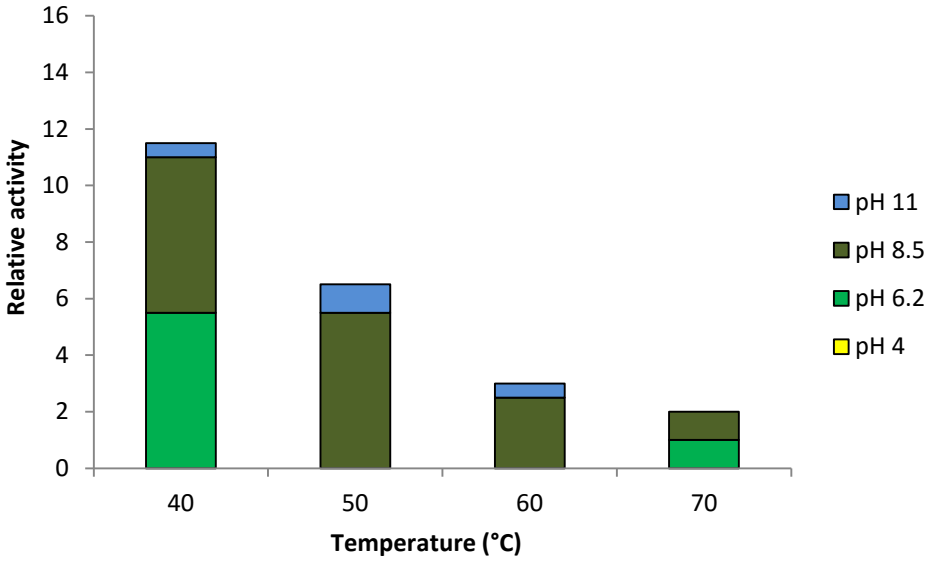


Figure 3.11 Potential *B. subtilis* AHU enzymatic activity across a pH and temperature range on xylan agar (n=2).

Isolate AHW was identified as *Brevibacillus agri* via 16S rRNA sequencing and was isolated at 50 °C, the gram stain and shape match this identification. Cultivation on xylan agar yielded only one point of positive enzymatic activity at pH 11 and 50 °C (Figure 3.12). However, the isolate was able to grow at 40-60 °C across the pH range,

from 6.2-11. AHW had no visible activity on starch agar though it was able to grow between pH 6.2 and 8 across the whole temperature range. AHW had very low growth and activity on CMC agar with an optimum of 40-50 °C.

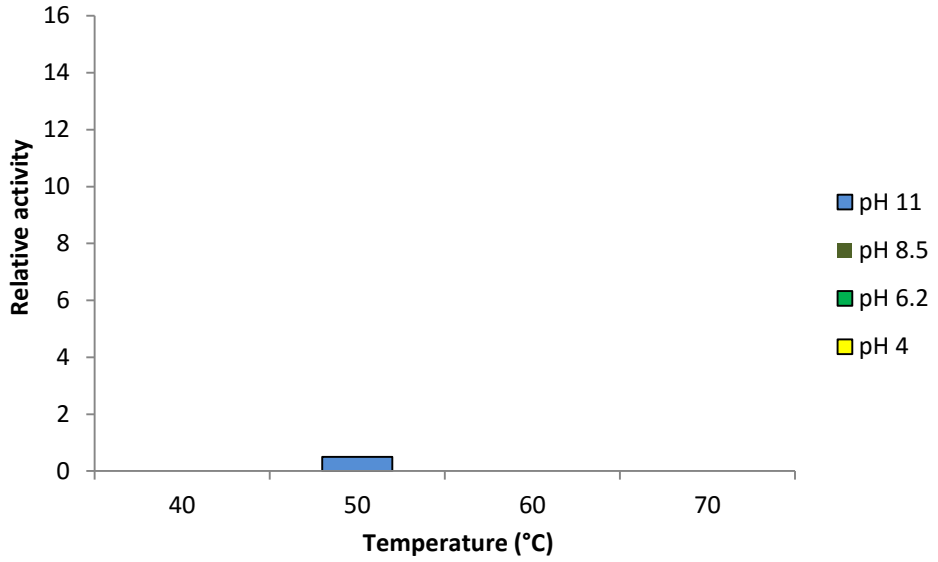


Figure 3.12 Relative enzymatic activity of potential *Brevibacillus agri* AHW across a pH and temperature range on xylan agar (n=2).

Isolate AHY1 was tentatively identified as *Bacillus licheniformis* (below 97 % identity on NCBI BLAST and RDP). However the Gram stain and shape match this identification (Holt *et al.*, 2000) and the bacterium was isolated at 50 °C which fits this characterisation. The optimal temperature of AHY1 correlates to this; however this isolate's peak pH appeared to be wider with much activity retained at pH 11 (Figure 3.13). There is no activity at pH 4 and little activity at 70 °C, which was reflected by bacterial growth; however there was limited growth at pH 4 at 40 °C. Interestingly there was little activity at pH 6.2 at 60 °C whereas this rose sharply at 50 °C. There was also a lot of activity at 50 °C at pH 11, however there was no activity at any other temperature at this pH.

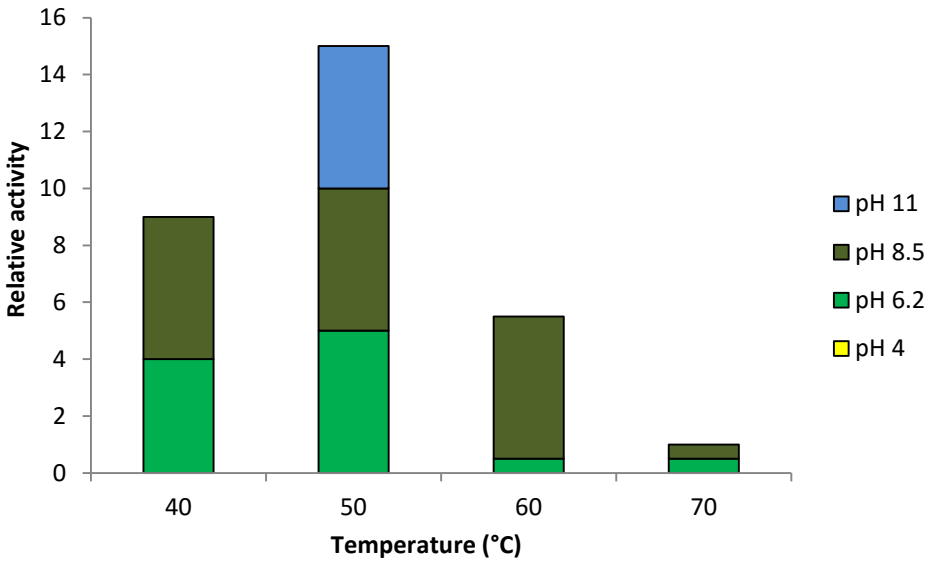


Figure 3.13 Potential *B. licheniformis* AHY1 enzymatic activity across a pH and temperature range on xylan agar (n=2).

Isolate AIB was also identified as *Bacillus licheniformis*; however the Gram stain appeared variable. This isolate showed similarity to AHY1 (also identified as *B. licheniformis*). There was no visible growth or activity at pH 4 at any temperature; however growth exceeded activity at most data points and though there was growth at pH 11 and 60 °C there was no visible activity (Figure 3.14). The majority of growth and enzyme activity appeared at pH 6.2 and 8.5 at the lower temperatures of 40 and 50 °C. Activity decreased the higher the temperature and at the extremes of pH. There was no visible growth detected at pH 8.5 at 70 °C; however enzymatic activity was low but was the same as pH 6.2 at the same temperature.

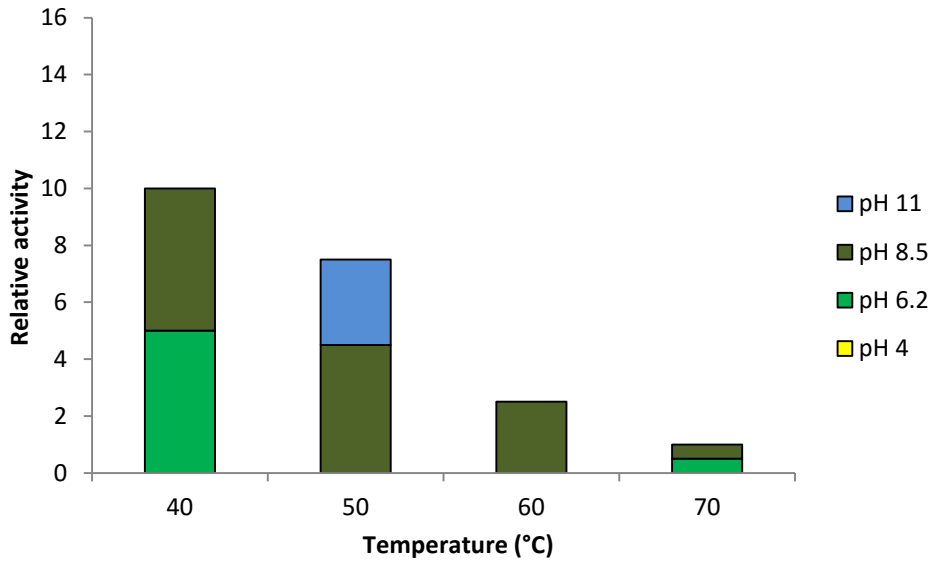


Figure 3.14 Relative enzymatic activity of potential *B. licheniformis* AIB across a pH and temperature range on xylan agar (n=2).

Isolate AIE1 was tentatively identified as *B. licheniformis* (below 97 % max identity on NCBI BLAST and RDP), though the Gram stain and shape match this identity (Holt *et al.*, 2000). Enzymatic activity was present across the temperature range with 6.2 as the optimal pH (Figure 3.15). However, there was high activity at pH 11 at 50 °C, though activity at pH 6.2 persisted across the whole temperature range.

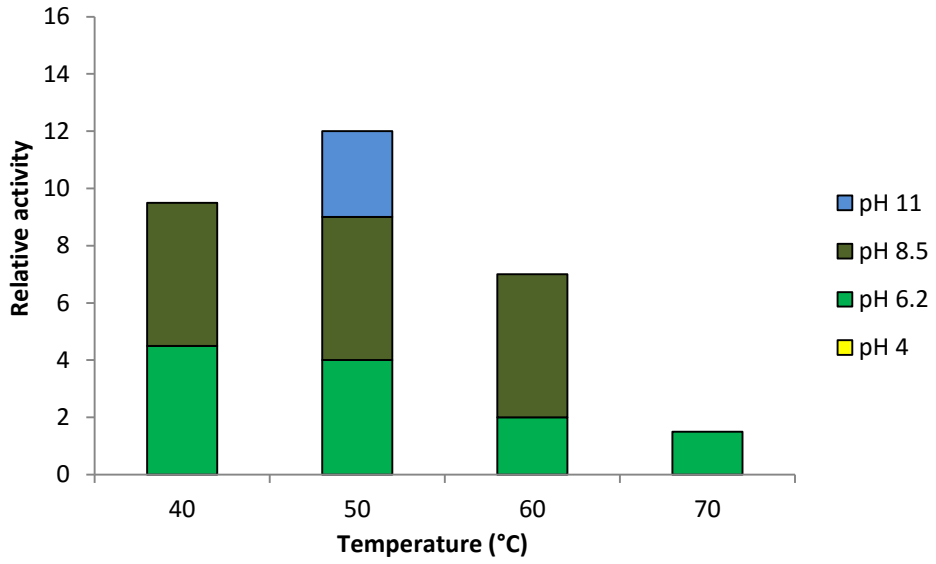


Figure 3.15 Potential *B. licheniformis* AIE1 enzymatic activity across a pH and temperature range on xylan agar (n=2).

AIF was identified as potentially a *Bacillus stratosphericus* and was isolated at 40 °C. There was no activity on either CMC or starch at any temperature or pH. However, there was considerable growth on the plates. Activity on xylan was low but consistent across the pH range at 40 °C, though an increase in temperature caused a decrease in activity which became none existent at 60 and 70 °C (Figure 3.16). This is however not unexpected due to the isolation temperature of this bacterium. The growth of AIF on xylan was considerably higher than the activity, and thus the enzymes used are most probably cell bound or within the cell.

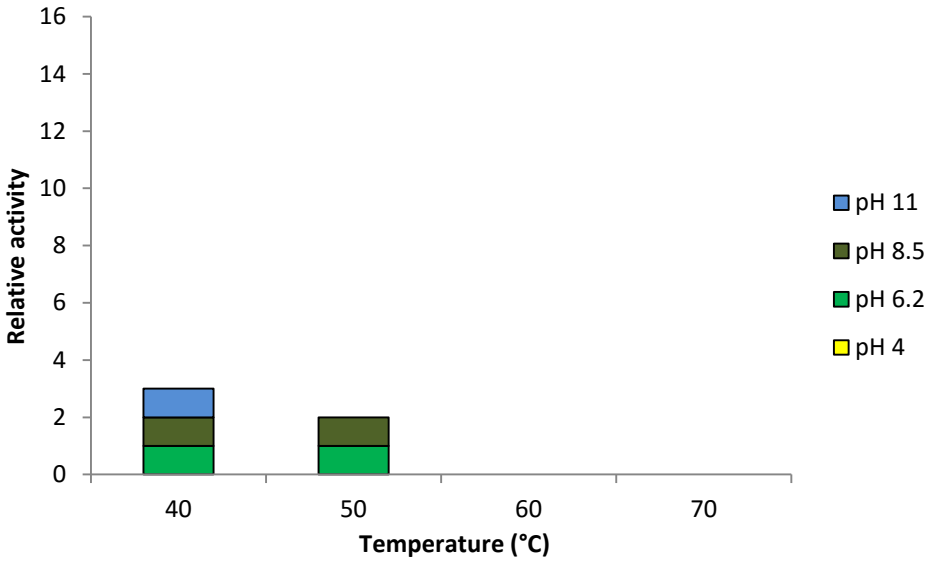


Figure 3.16 Relative enzymatic activity of potential *B. stratospericus* AIF across a pH and temperature range on xylan agar (n=2).

AIK was also identified as *B. licheniformis*, with Gram stain and shape matching this ID (Holt *et al.*, 2000), and had little or no activity on CMC and starch. Activity on xylan agar peaked at pH 11 at 40 °C, though growth was most prevalent at pH 6.2 at the same temperature (Figure 3.17). Activity and growth on xylan, demonstrate that though growth was high between 40 and 50 °C, activity was proportionately lower and

did not extend to 60 °C. Growth peaked at pH 6.2 and 40 °C whereas activity was highest at pH 11 at the same temperature.

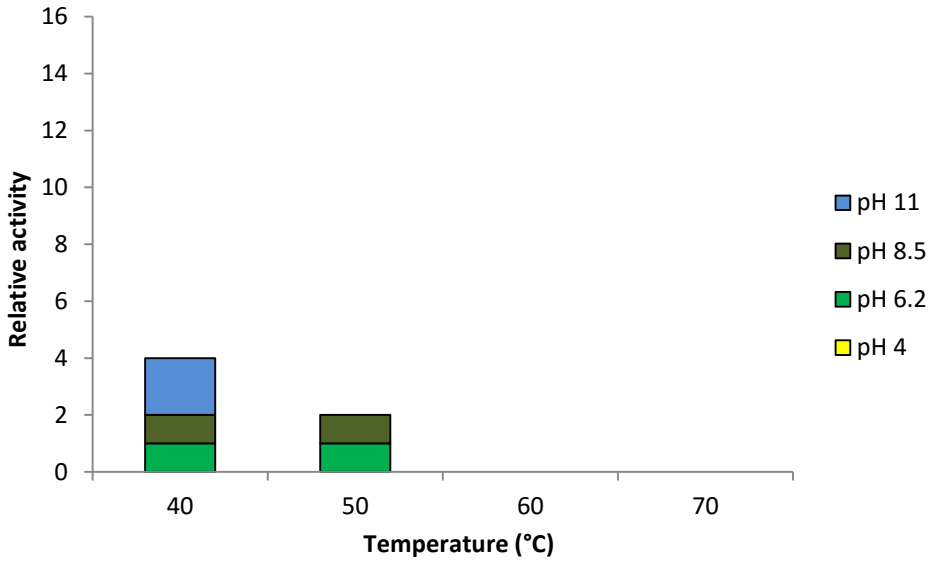


Figure 3.17 Potential *B. licheniformis* AIK enzymatic activity across a pH and temperature range on xylan agar (n=2).

AJF1 was identified via 16S rRNA sequencing, plate morphology and Gram staining as *Bacillus licheniformis* (Holt *et al.*, 2000), and was isolated at 40 °C. Again this isolate had high growth and activity across the temperature and pH range bar pH 4 (Figure 3.18). Activity and growth on xylan agar is very similar to CMC agar with the peak for both being pH 8.5 at 40-50 °C. However there was activity at 50 °C at pH 11 corresponding with growth which spanned 40-60 °C at this pH. There was also activity, and growth to a higher proportion, at 70 °C but this only occurred at pH 6.2.

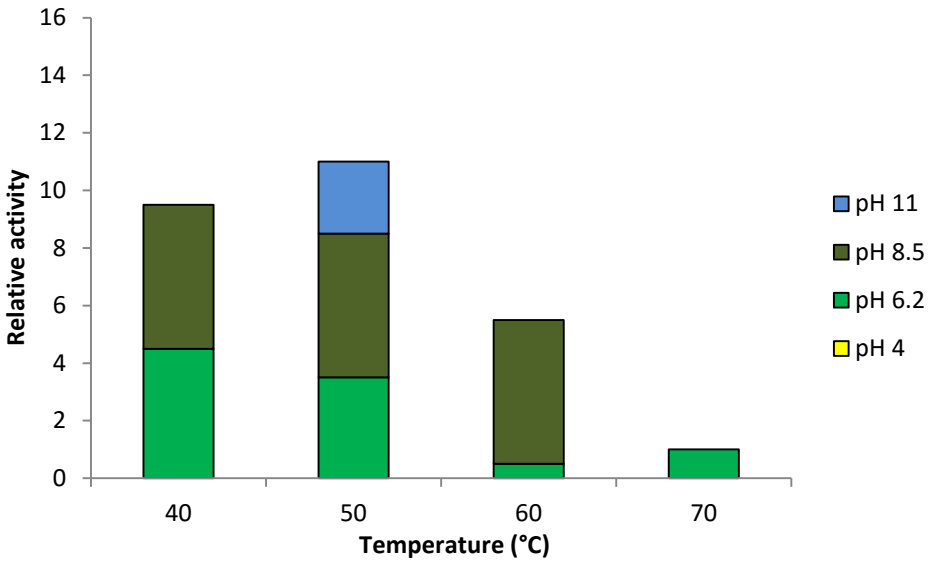


Figure 3.18 Relative enzymatic activity of potential *B. licheniformis* AJF1 across a pH and temperature range on xylan agar (n=2).

AJF2 was identified as a *Bacillus subtilis* using 16S rRNA and Gram staining (Holt *et al.*, 2000) and was isolated at 40 °C on MRS agar (though it could grow happily on Nutrient agar also). Both growth and activity were proportionately higher on xylan agar than CMC. Here activity peaked at pH 8.5 and 40 °C with growth peaking at pH 6.2 and 50 °C. Activity and growth were both visible at 70 °C at pH 6.2 though there was a drop in both growth and activity at 50 °C at this pH. Activity and growth on xylan at pH 11 both peaked at 50 °C but was also present at 40 and 60 °C.

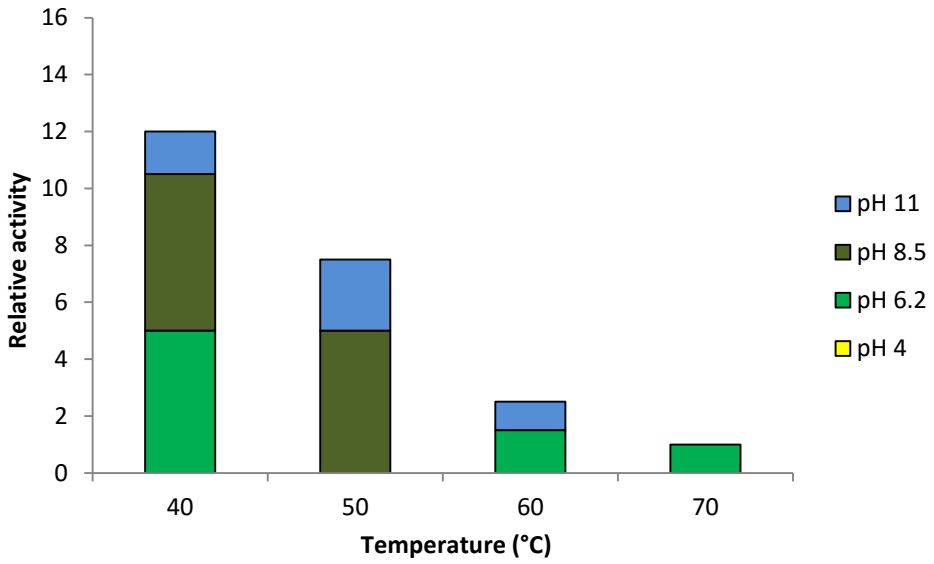


Figure 3.19 Potential *B. subtilis* AJF2 enzymatic activity across a pH and temperature range on xylan agar (n=2).

AJG was identified as potentially *Bacillus licheniformis* and was isolated at 40 °C on MRS agar, though it was able to grow on Nutrient agar also. As with all other bacterial isolates, AJG was identified using 16S rRNA sequencing, Gram staining and microscopy (Holt *et al.*, 2000). Activity on xylan agar was proportionately higher than growth meaning that enzymes were secreted into the media, with peak activity at pH 8.5 and 50 °C (Figure 3.20). There was also activity on pH 11 agar from 40-60 °C though no activity at 70 °C at the most alkaline pH. However there was activity at 70 °C on pH 6.2 media though no observable growth.

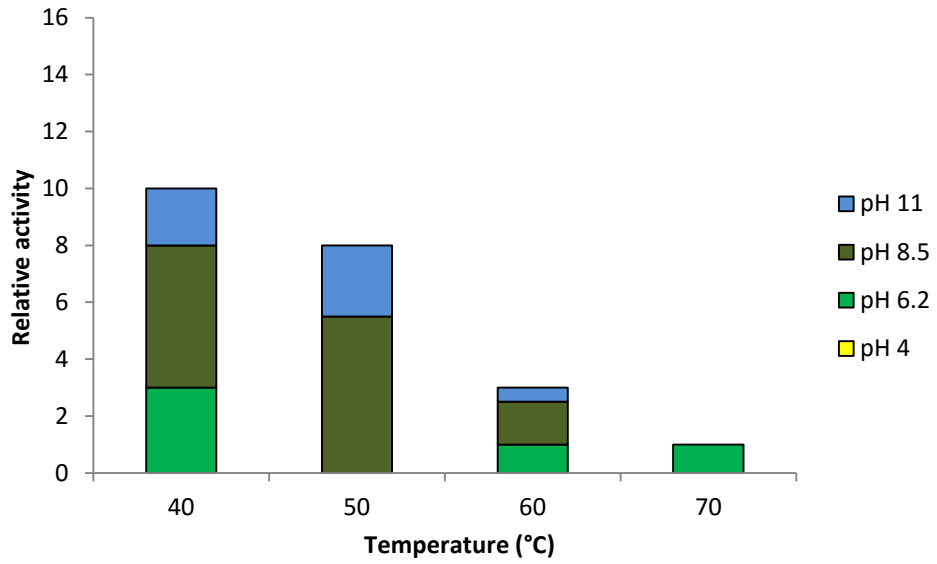


Figure 3.20 Relative enzymatic activity of potential *B. licheniformis* AJG across a pH and temperature range on xylan agar (n=2).

AJH1 was identified as a *Bacillus subtilis* via 16S rRNA sequencing and Gram staining (Holt *et al.*, 2000). This bacterium was isolated at 40 °C on MRS agar though it was able to grow on Nutrient agar also. Activity and growth of this isolate on xylan agar were very similar to on cellulose. Here activity at the extremes was present at 40 and 50 °C at pH 11, and pH 6.2 and 8.5 at 70 °C, though growth was only observed at pH 8.5 for the latter (Figure 3.21). Interestingly, growth was proportionately higher at pH 11 than activity at this pH. Peak activity stretched across 40-60 °C at pH 8.5 with peak growth appearing at pH 11 and 50 °C remarkably. Again enzymes were most probably secreted into the media as the halos of degradation were larger than the amount of bacterial growth present.

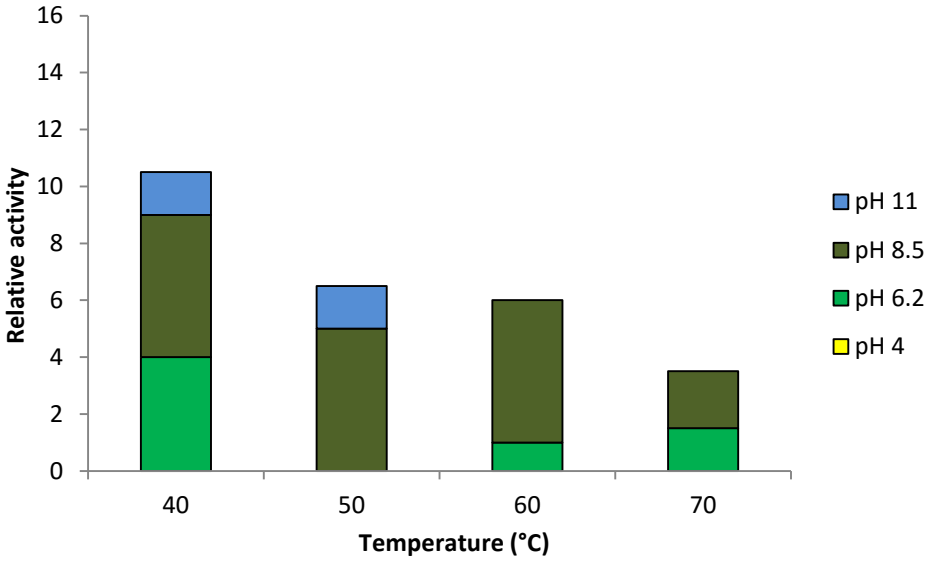


Figure 3.21 Potential *B. subtilis* AJH1 enzymatic activity across a pH and temperature range on xylan agar (n=2).

Isolate AJK was identified as potentially a *B. licheniformis* with the Gram stain and shape matching this 16S rRNA ID (Holt *et al.*, 2000). AJK was isolated at 40 °C overnight on MRS agar, but could be grown on Nutrient agar also. Activity on xylan agar was more widespread across the pH range than on CMC. Growth however did not seem to follow the same pattern as activity, whereby there was a large decrease in bacterial growth at 60 °C. However for both growth and activity 50 °C is the preferred temperature, with peaks at pH 8.5 for activity and pH 6.2 for growth (Figure 3.22). There was also activity at pH 11 at 50 °C as well as growth from pH 6.2 to 11 at this temperature and growth from 40 to 60 °C at pH 11. There was activity at 70 °C at pH 6.2, but growth across the pH 6.2 and 8.5 at this temperature.

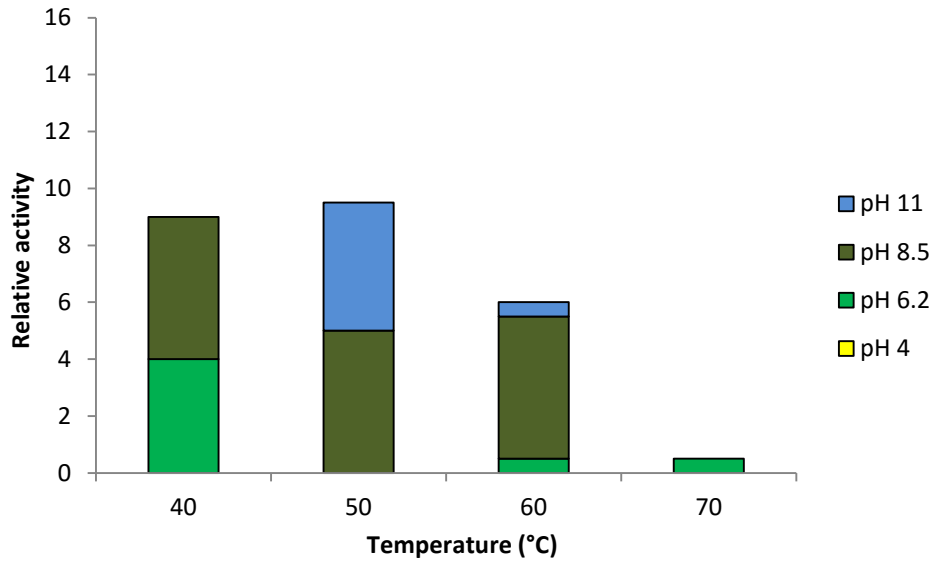


Figure 3.22 Relative enzymatic activity of potential *B. licheniformis* AJK across a pH and temperature range on xylan agar (n=2).

Isolate AJN1a was identified via 16S rRNA sequencing to be *Bacillus licheniformis*, however the Gram stain appeared to be negative rather than positively stained, which did not match the 16S ID (Holt *et al.*, 2000). AJN1a was isolated at 40 °C overnight on nutrient agar and was observed to be small cigar shaped rods. This bacterial isolate had activity across the carbon source range especially on CMC and xylan. Activity on xylan agar was highest at pH 6.2 at 40 °C, though activity was still high at the same temperature at pH 8.5 and 50 °C at pH 8.5 (Figure 3.23). Activity was sustained across the temperature range at pH 8.5 though activity was slightly higher at 70 °C at pH 6.2. However activity at pH 6.2 was not consistent as it was with pH 8.5. There was no activity or growth at pH 4 across the temperature range; however there was activity at 40 and 50 °C at pH 11 though growth was observed at 40 and 60 °C at this pH. Most growth appeared at pH 6.2 at 50 °C which dropped at 60 °C then increased again at 70

°C. No growth could be observed at pH 8.5 at 70 °C contrary to a halo of activity being observed.

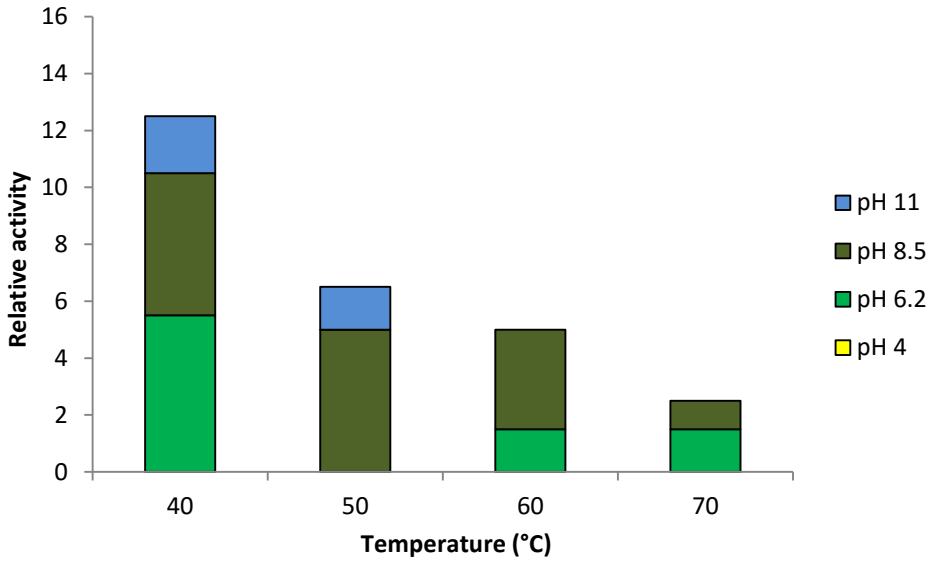


Figure 3.23 Potential *B. licheniformis* AJN1a enzymatic activity across a pH and temperature range on xylan agar (n=2).

AJN1b was identified as a *Bacillus licheniformis* via 16S rRNA sequencing, Gram staining and microscopy, and was isolated at 40 °C overnight on nutrient agar. Again activity and growth followed the same pattern when AJN1a was inoculated onto xylan agar at a range of pH and temperatures (Figure 3.24). However unlike on CMC agar peak activity and growth was observed at pH 6.2 and 40 °C. Growth could be seen at pH 6.2 at 70 °C but activity was apparent at both pH 6.2 and 8.5 at this temperature. Plus although growth was observed at 60 °C at pH 11, activity could be seen at 40 and 50 °C but not at 60 °C. Again there was no growth or activity at pH 4 across the temperature range.

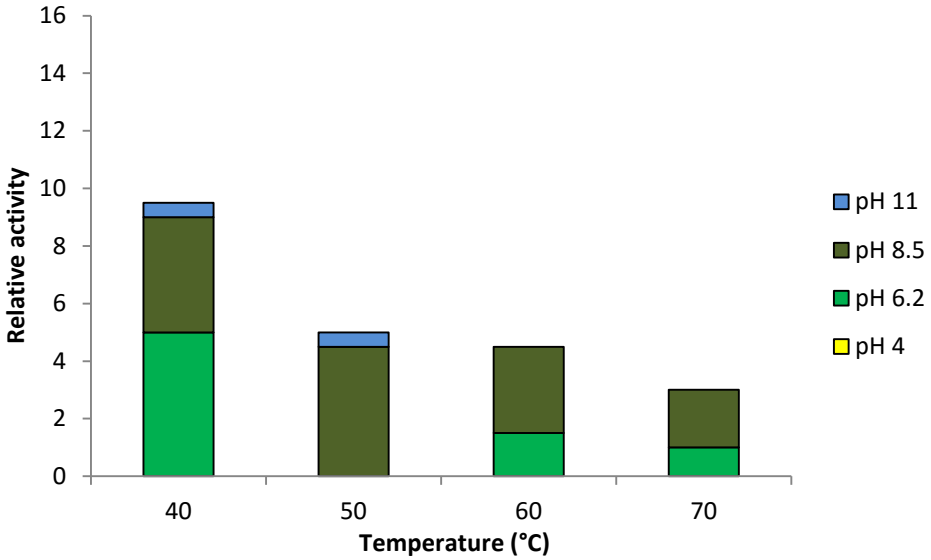


Figure 3.24 Relative enzymatic activity of potential *B. licheniformis* AJN1b across a pH and temperature range on xylan agar (n=2).

Isolate AJT was putatively identified as *B. licheniformis* though the 16S rRNA sequencing hits on NCBI BLAST and RDP had below 97 % maximum identity, however the Gram staining and microscopy agreed with this ID. The activity and growth of isolate AJT peak at pH 8.5 at 40 °C on xylan agar (Figure 3.25). Though growth was observed from pH 6.2-11 at 70 °C, and 40-70 °C at pH 11 activity was only recorded at pH 6.2 at 70 °C, and 50 °C at pH 11. There was a dramatic drop in activity and growth at 70 °C compared to the previous temperatures across the pH range.

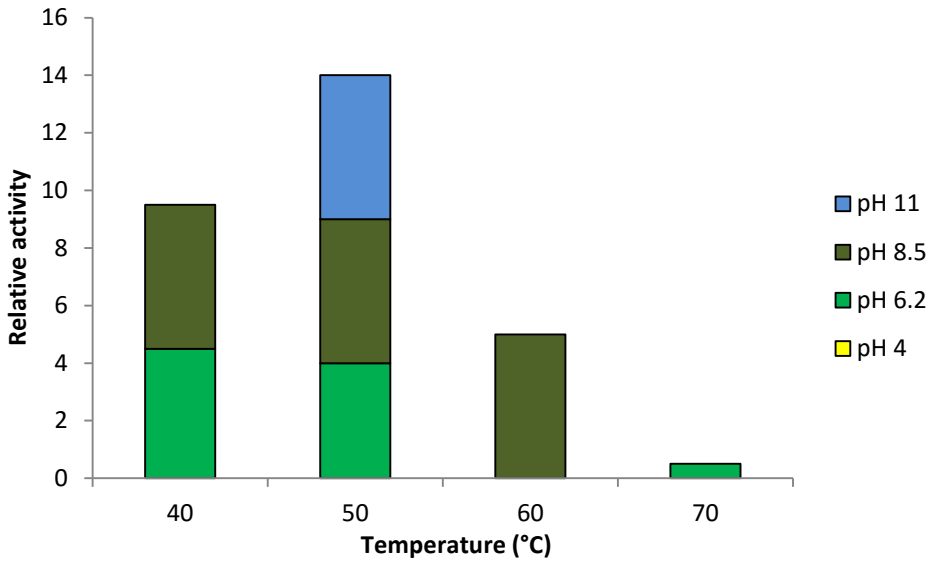


Figure 3.25 Potential *B. licheniformis* AJT enzymatic activity across a pH and temperature range on xylan agar (n=2).

AJVa was identified via 16S rRNA sequencing as *B. licheniformis*, however the Gram stain was negative and did therefore not match the *Bacillus* genus (Holt *et al.*, 2000). This bacterium was isolated at 50 °C overnight on Nutrient agar and was also subject to the carbon utilisation assays. On xylan agar activity is proportionately higher than growth with peaks at 40-50 °C at pH 8.5 (Figure 3.26). Interestingly there was growth across the temperature range at pH 11 though only activity at 50 °C. There was also activity at pH 6.2 at 70 °C though no growth was detected.

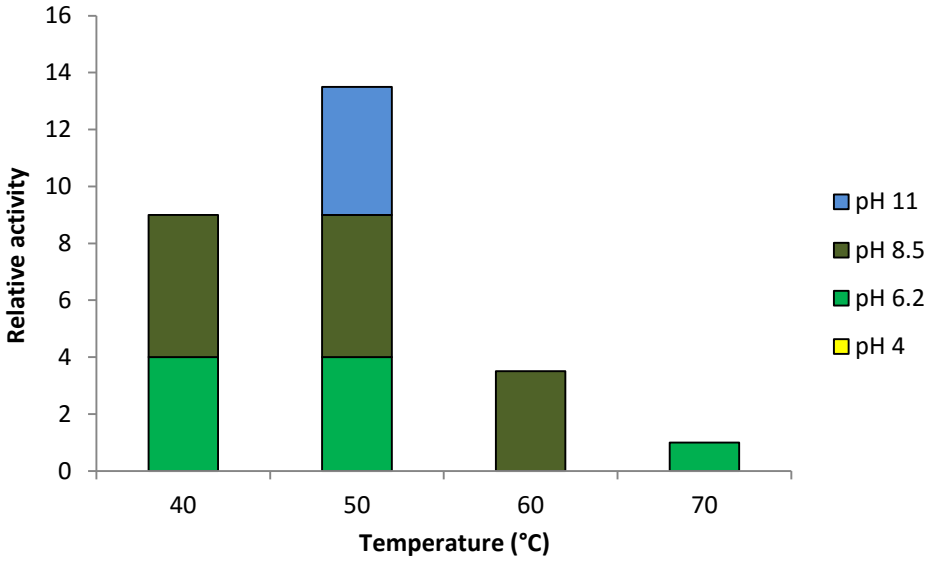


Figure 3.26 Relative enzymatic activity of potential *B. licheniformis* AJVa across a pH and temperature range on xylan agar (n=2).

3.3 Discussion

Aerobic bacteria were isolated from degrading *Miscanthus* chip and soil surrounding *Miscanthus* plots in order to identify isolates that had potential enzymatic activity, on cell wall components, for potential utilisation on pre-treated biomass. Initially, gDNA was isolated from soil and compost, with the aim to generate fosmid libraries. However, a culture based method was adopted to find aerobic isolates from decomposing *Miscanthus* chip. Such a method allowed complete analysis of each isolate, avoiding activities potentially missed when using fosmid libraries. Isolates were hypothesised to have saccharolytic abilities from such sampling due to the isolation environment (P. Gupta, Samant, & Sahu, 2012; Ko *et al.*, 2007; F. B. Yu, Shan, Luo, Guan, & Qin, 2013). Initial isolation took place on Nutrient agar, as it was an all-purpose media (Thermo Fisher Scientific Inc., 2014), and would therefore allow growth of a greater range of bacteria. MRS media was used for a directed isolation approach

(lactic acid bacteria specific media). Lactic acid had been produced from xylose (Okano, Tanaka, Ogino, Fukuda, & Kondo, 2010) and cellulose (Yáñez, Moldes, Alonso, & Parajó, 2003); thus lactic acid media was an initial step in directed selection. Thermotolerant isolation utilised both media types for selection; morphology was drastically different on both media. Single colony isolation typically took place from Nutrient agar plates, due to high growth rate and potential extracellular polysaccharide production on MRS agar due to high carbohydrate content (Cieslewicz, Kasper, Wang, & Wessels, 2001; Donkersloot, de Leon, Chassy, & Krichevsky, 1976; Ruas-Madiedo & de los Reyes-Gavilán, 2005). Mass spectrometry could have been used to check the polysaccharide content (Kind *et al.*, 2012).

Bacterial identification using the 16S rRNA hypervariable regions was used due to ease and success rate of this technique (Janda & Abbott, 2007). Sequence data were analysed using NCBI BLAST and RDP BLAST, with the lowest e-values and highest maximum identity (above 97%) being taken for identification (Petti, 2007). However, in some instances the two databases reported confounding positive identification (see Appendix 1, section 8.1.2). A Neighbour-Joining (NJ) phylogenetic tree was used to display all 16S rRNA sequences (Figure 3.2) with the potential identification from BLAST (Appendix 8.1.2) preceding the unique three letter code of each isolate. The NJ method was used due to the speed of computation and accuracy with smaller data sets (Tamura *et al.*, 2004). The tree split into seven clades whereby the most closely related isolates grouped together. Of the five main clades (excluding mixed clades 1 and 2, Figure 3.2) four were identified as Gram negative via sequence homology. Gram positive isolates were grouped in the *Bacillus sp.* clade or scattered into the two mixed

clades (*Microbacterium sp.*, *Micrococcus luteus* ADM, *Streptomyces sp.* and *Staphylococcus sp.*). However, Gram staining did not reflect the 16S sequencing results, as the majority of isolates were Gram positive from Gram stain analysis (Figure 3.3). Gram staining and 16S rRNA identification may not have agreed due to human error in Gram stain analysis or due to aging of cells. *Bacillus sp.*, *Butyrivibrio sp.* and *Clostridium sp.* have all been demonstrated to change to Gram negative as the culture aged and cells became more fragile (Beveridge, 1990). As isolates grew at different rates it could be possible that some were in differing growth stages to others when Gram stained.

Isolates found within the five main clades have been reported to have a number of applications (Table 8.1). A number of *Burkholderia spp.* were reported to be endophytic, promote plant growth (Mohanta, Sharma, & Deb, 2010), pathogenic to humans and plants (Du Toit, 2015), and hydrolyse hemicellulose in biorefining (FitzPatrick, Champagne, Cunningham, & Whitney, 2010). Some *Luteibacter spp.* were reported to be symbiotic and promote growth in plants, as well as be of use in biotechnology, via xylose and cellobiose utilisation (Li Wang, Wang, Li, & Jiang, 2011); bio-fertilisation of soils (Guglielmetti *et al.*, 2013) and degrade pollutants (Bresciani *et al.*, 2013). Some *Chryseobacterium spp.* were reported to be potential human pathogens (Berg, Eberl, & Hartmann, 2005; Hoque, Graham, Kaufmann, & Tabaqchali, 2001), as well as have potential biorefining uses (S. L. Wang, Liang, & Liang, 2011). However, both *Chryseobacterium* and *Luteibacter* were not recorded as dangerous in The Approved List of Biological Agents (Advisory Committee on Dangerous Pathogens, 2013). *Pseudomonas spp.* were reported to have biotechnological uses via

fermentation (Z. H. Xu, Bai, Xu, Shi, & Tao, 2005), enzyme production (Raghothama *et al.*, 2000; Shah & Gupta, 2008), and biodegradation of crude oil (Rahman, Thahira-Rahman, Lakshmanaperumalsamy, & Banat, 2002) and herbicide (Herrera-González *et al.*, 2013). However *P. aeruginosa* was reported as the most common pathogen to infect cystic fibrosis patients and is a known human pathogen (Govan & Deretic, 1996). *Bacillus spp.* are well known for their biotechnological uses, especially in enzyme production (A. de Boer *et al.*, 1994; Rhee *et al.*, 2011; Schallmey *et al.*, 2004); they have also been used to prevent soil-borne plant diseases (Gavrilescu & Chisti, 2005). Thus, though many isolates were recorded as potential human pathogens, there was also the potential for isolation of useful novel enzymes. All species had to be taken into consideration as many isolates were only identifiable down to genus level (Figure 3.1).

Where isolates were mixed (i.e. there were two or more different bacteria present) a range of primers were employed for identification to try and avoid bias (Polz & Cavanaugh, 1998; Suzuki & Giovannoni, 1996). Serial dilutions (Janssen, Yates, Grinton, Taylor, & Sait, 2002; Schut *et al.*, 1993) and pour plates were adopted to separate mixed colonies, however 19 cultures remained mixed throughout. Pour plates damage temperature sensitive cultures (Hoben & Somasegaran, 1982; Van Soestbergen & Lee, 1969) and were therefore not investigated in great depth. Morphology of single colony cultures on agar plates, including form, elevation and margin, were compared to descriptions in the literature to aid Gram stain analysis and taxonomic identification.

The bacterial library mainly consisted of rod shaped isolates (76.9% in mesophiles and 88% in thermotolerants; section Chapter 3). There were proportionally more dry

surfaced strains in the thermophile collection (11%) than the mesophilic collection (5.8%). However, morphology was described based on thermophiles cultured on Nutrient agar; when grown on MRS agar morphology was drastically different (Appendix 8.1.6). Morphology form was more diverse in the mesophilic isolate collection, where variance of species was larger. The decreased range in bacteria isolated at higher temperatures was due to the temperature range. Fewer species were able to grow at elevated temperatures (Stolp & Starr, 1981).

Mixed cultures, that could not be separated and properly identified were not discarded, as undefined mixed cultures were shown to be biotechnologically useful (Agler, Wrenn, Zinder, & Angenent, 2011). Additionally, biodegradation of lignocellulose most efficiently occurs in the guts of insects and some mammalian species, due to highly evolved mixed cultures (Angenent & Wrenn, 2008). Therefore, mixed cultures are able to withstand complex and variable substrates due to metabolic flexibility, owing to the changeability and diversity in the community (Angenent & Wrenn, 2008; Kongjan, O-Thong, Kotay, Min, & Angelidaki, 2010). Anaerobic, undefined mixed cultures for use in the carboxylate platform, have also been shown to render sterilization and aeration superfluous, thus reducing costs (Agler *et al.*, 2011). Mixed colonies could therefore aid carbon utilisation (Agler *et al.*, 2011; Angenent & Wrenn, 2008; Halsall & Gibson, 1985). Thermotolerant mixed cultures have been shown to produce large amounts of hydrogen for use as biofuel (Kongjan *et al.*, 2010), and a mixed mesophilic bacterial consortium has been shown to degrade crude oil contamination (Rahman *et al.*, 2002). However, products of interest are sometimes

converted to other compounds faster than they are formed, during fermentation with mixed cultures (Angenent & Wrenn, 2008).

The most frequently occurring isolates in the bacterial library were *Burkholderia fungorum*, *Pseudomonas sp*., *Bacillus licheniformis*, and *Luteibacter rhizovicinus*. All five species have been shown to be soil and/or plant associated (Berg *et al.*, 2005; Bresciani *et al.*, 2013; Glick, 2012; Johansen, Binnerup, Kroer, & Mølbak, 2005; Kirchhof, Reis, Baldani, Eckert, & Hartmann, 1997; J. Li, Dai, Chen, & Zhu, 2007; Rey *et al.*, 2004; Uffen, 1997). Though *Burkholderia sp.* have been shown to be efficient at degradation, in a biorefining context (FitzPatrick *et al.*, 2010), and tolerant of organic solvents (Ferrer *et al.*, 2007), the genus has been rejected by the US Environmental Protection Agency due to the many associated disadvantages. These include deadly infection in cystic fibrosis patients; plant and domestic animal pathogen, and food poisoning agent (Andreolli, Lampis, Zenaro, Salkinoja-Salonen, & Vallini, 2011). British regulations class *B. cepacia* as hazard group 2, and *B. mallei* and *B. pseudomallei* as hazard group 3 (Advisory Committee on Dangerous Pathogens, 2013). However, some *Burkholderia* species have been classed as gram-negative non-fermenters (Vaneechoutte, Dijkshoorn, Nemec, Kämpfer, & Wauters, 2011) and would therefore be useful as a source of enzymes. Some *Pseudomonas sp*. have also been classed as hazard group 2 (*P. aeruginosa*) (Advisory Committee on Dangerous Pathogens, 2013). However, six *Psuedomonas sp.* have been shown to utilise xylose, though a number of species within this genus are known human, animal and plant pathogens (Holt *et al.*, 2000).

In the bacterial library 29 *Bacillus sp.* were isolated and identified including 13 *B. licheniformis* isolates. A plethora of literature details the enzymatic capabilities of *B. licheniformis*, especially for alpha-amylase isolation due to its stability over a pH range (Hmidet *et al.*, 2008; Ramesh & Lonsane, 1989) and temperature (Declerck *et al.*, 2003; Declerck, Machius, Wiegand, Huber, & Gaillardin, 2000). More recently, attention has turned to cellulolytic and hemicellulolytic enzymes, and thus *B. licheniformis* has been investigated (van Dyk, Sakka, Sakka, & Pletschke, 2009, 2010). The *Bacillus* species are widely recognised as a soil and rhizobium related bacteria, though they have been isolated from a wide range of sources. They are therefore suitable for investigation of their plant material degrading abilities (Encylopaedia Britannica Inc., 2014; Kenyon College, 2010; van Dijl & Hecker, 2013). *B. licheniformis* and *B. amyloliquefaciens* are two bacteria that are most frequently used for commercial uses, as they have amylases that meet industrial demands (Hmidet *et al.*, 2008). In the bacterial library, 13 strains isolated were identified as *B. licheniformis*. These isolates were highly active on carbon utilisation assays (section 3.2.5.2), all bar AIK were active on at least one carbon source at 70 °C. AIK was later identified as most likely to be *Bacillus pumilus* (section 5.3).

In the bacterial collection 13 *Luteibacter sp.* were identified. *Luteibacter rhizovicinus* has been described as Gram negative rods, yellow pigmented, aerobic, motile (polar flagella) and chemo-organotrophic (Johansen *et al.*, 2005). Isolates that were identified as *Luteibacter sp*. (AEZa and AEZb) or *L. rhizovicinus* (AEOa, AFBa and AFBb) were cream coloured (Appendix 1: 8.1.2). However, four of these were mixed colonies with both Gram positive and Gram negative rods/cocci. The other bacterium (AEZa)

was Gram variable and thus also had both positive and negative rods. American Type Culture Collection (ATCC) states that *L. rhizovicinus* is yellow, tiny, circular, smooth, entire and translucent (ATCC, 2013), which was very close to the morphology recorded for many of the isolates identified as this bacterium in 8.1.2. However, *Dyella sp*. have also been described as yellow, rod shaped, motile bacteria (Jung *et al.*, 2009) and had high similarity scores to many of the bacteria identified as *Luteibacter sp*.. *L. rhizovicinus* has been shown to have environmental application in the degradation of herbicide propanil and its break down products to reduce toxicity in soil and water (Herrera-González *et al.*, 2013), as well as association with Pine Wilt disease (Proença *et al.*, 2010). *Luteibacter* have little association with biotechnology or biorefining bar bioremediation of oil refinery sludge (J. Zhang, Li, Chen, & Thring, 2011) and extracellular lipase production (Bresciani *et al.*, 2013). Potential for novel enzyme discovery is therefore higher.

Within the bacterial collection 42.3% of mesophiles were Gram positive whereas 73% of thermotolerant isolates were Gram positive. However, there were a higher proportion of mixed colonies within the mesophilic group (17.3%). Nevertheless, 35.8% of isolates had either incorrect 16S identification or Gram stain, which could be due to the presence of mixed cultures. Certain *Bacillus sp.* (including *B. subtilis* and *B. licheniformis*) as well as *Moraxella sp.* were shown to give undefinable readings from Gram staining. *Moraxella sp.* appear Gram positive when they are negative (Gregersen, 1978), as demonstrated by isolates ABF2 and ADK, which were identified via 16S as possible *Enhyrdobacter sp.* or *Moraxella sp.*. Treatment with potassium hydroxide (KOH) was reported to aid Gram staining. Gram negative cells lyse after 1

min of KOH treatment (Gregersen, 1978; Halebian, Harris, Finegold, & Rolfe, 1981; Suslow, Schroth, & Isaka, 1982), whereas *Bacillus* strains were still intact after 30 min treatment (Gregersen, 1978).

Growth curves were conducted to aid morphological assessment as ABR grouped with *Pseudomonas sp*. ACO (Figure 3.2). *Bacillus coagulans* ABP and ABQ had similar growth curves, whereas ABR had a longer lag phase before heading into a slow exponential phase (Figure 3.6). *B. coagulans* ABR exponential phase was eventually much steeper than ABQ, but nearly parallel to ABP. *B. coagulans* ABP has the highest optical density and thus reached a higher amount of biomass before ABQ and ABR. After 25 hours growth all three isolates reach the same optical density. It is unknown why *B. coagulans* ABP and ABQ had a drop in cell density before the exponential phase, which appeared later for ABR. *B. coagulans* ABP and ABQ followed the expected growth curve shape (Khurshed, 2003). The pH also dropped by 1.81 (ABP), 1.72 (ABQ) and 1.8 (ABR) after 25 hours of growth (starting media was pH 6.2). A build-up of acid metabolic products had probably occurred. HPLC analysis could be used to verify this.

The carbon utilisation assays were developed as a crude test of enzyme activities of each of the isolates. All bacterial isolates were tested in duplicate across a pH range and their results averaged (Appendix 8.1.3). Isolates chosen for further characterisation were shown in Figure 3.8. Thermotolerant isolates were tested across the pH range and a temperature range simultaneously (Figure 3.11 to Figure 3.26). These data demonstrated that isolates were active across the carbon source range; active across the pH range; active only at one pH on certain carbon sources, or had no activity at all. Activity at low pH was minimal, whereas at near neutral pH activity was

high. Only eight mesophilic isolates had activity at pH 4 on xylan agar (Figure 3.8), though there was no activity on CMC at this pH. However, 27 mesophilic isolates had activity at pH 4 on starch agar (8.1.3). Of the seven isolates chosen for whole genome sequencing only two (AEH1 and ABQ) had no activity at pH 4 on starch agar. When tested for alkaline tolerance on CMC (pH 10) and xylan (pH 11), there were no positive results for the mesophilic isolates, though 15 were active at pH 9 on starch agar. Isolates were chosen for further characterisation by whole genome sequencing (section 5.2) based on their carbon utilisation profile. These mesophilic bacterial isolates were all active at pH 4. Therefore, such isolates probably produce enzymes that were acid tolerant, which could be useful for ethanol, acetone and butanol fermentation (Lowe, Jain, & Zeikus, 1993). Many bacterial species have an internal pH of near neutral (pH 7). Those that have adapted to more extreme conditions, are able to alter their internal pH thereby avoiding an inverted membrane potential (Siegumfeldt, Rechinger, & Jakobsen, 1999). To prevent cytosol acidification by H^+ (Schäfer, Engelhard, & Müller, 1999) from the environment, the cell membrane is maintained by an electrochemical gradient of protons, and movement of alkali cations K^+ and Na^+ across the cytoplasmic membrane (Bakker, 1990; Michels & Bakker, 1985). Bacterial isolates with carbon utilisation profiles of activity at pH 4, could be using such proton motive force to survive in an acidic environment. As mesophilic isolates, these bacteria or enzymes could be used for initial saccharification of biomass (after pre-treatment to remove lignin) due to their preference for low temperatures. Yeast fermentations could then be utilised, without additional costs of cooling (Abdel-Banat *et al.*, 2010; La Grange *et al.*, 2001; Leandro *et al.*, 2006; Toivari *et al.*, 2004;

Walfridsson *et al.*, 1996), as the optimal temperature for xylanase enzymes is 50 °C (Biocatalysts Ltd., 2015a, 2015b; Novozymes Bioenergy, 2010b).

There are two classes of enzyme distinguished by their differing endo- and exo-active sites (Gerardi, 2003). Exo-active enzymes have a tunnel active site (Claeyssens & Henrissat, 1992; Teeri, 1996; van Aalten *et al.*, 2001), though some reports of microbes in the literature contradict this structural claim such as *Humicola insolens'* cellobiohydrolase (Varrot, Schulein, & Davies, 1999). This mode of action cuts chains of sugars from the end, liberating glucose/fructose monomers (Fujii, Murakami, Yamada, Ona, & Nakamura, 1981), but doing little to alter protein molecular weight compared to endo-enzymes (Wheatley & Moo-Young, 1977). Such action could explain the faint and diffused halos present on some of the plates (Figure 3.7, Figure 3.9 and Figure 3.10), as chains are being degraded more slowly in length and thus have more binding positions for the iodine (S. E. Lee, Armiger, Watteeuw, & Humphrey, 1978; Suga, Dedkm, & Moo-Young, 1975). Endo-active enzymes cut molecules at random sites splitting them into two (Fujii *et al.*, 1981), and have an active site cleft allowing them greater access to the substrate molecule (Harjunpää, Helin, Koivula, Siika-aho, & Drakenberg, 1999; Horn *et al.*, 2006; Torronen & Rouvinen, 1995). Such enzymes could produce the clear halos demonstrated on CMC in Figure 3.7. Extracellular enzymes, which are secreted and thus produce larger halos on the agar, were preferred to intracellular enzymes (that are within the cell), as they are more likely to be tolerant to heat and pH (Bragger *et al.*, 1989). Such halos were demonstrated on CMC and xylan agar (Figure 3.7); however halos on starch agar were often only present where growth had been visible. For each assay, iodine was used to dye the agar plates, because it

binds to long chain sugars (Morrison & Karkalas, 1990; Williams, 1983). When broken down, clearance zones were formed either under or around each colony (Kasana, Salwan, Dhar, Dutt, & Gulati, 2008; Ko *et al.*, 2007; Pointing, 1999). However, in some cases (AIF and AIH1), growth was present but no halo. Therefore, isolates were able to grow on the basal medium without utilising the carbon source or very limited carbon source. Yeast extract could have been used as a carbon source (Melford, 2011; Neogen Corporation, 2011), though it was at a low concentration (0.1 g/L). Alternatively, isolates may have been able to utilise carbon from basal media components, such as diAmmonium tartrate, as *Pseudomonas sp.* have been shown to be able to survive on C1 compounds (Anthony, 1975; Peel & Quayle, 1961). However, there is no relevant literature to validate these claims.

Most of the isolates chosen for whole genome sequencing (section 5.2) were thermotolerant strains. Isolates were selected due to their activity at alkaline pH levels, as well as their prevalence at high temperatures. Isolates growth tended to correspond with activity, and growth was generally higher than activity on xylan agar. This could be due to the enzyme remaining in the cytosol rather than being secreted, via a signal peptide which causes it to be transported to the cell wall, and excreted into the agar (Honda & Kitaoka, 2004; Martoglio & Dobberstein, 1998). AHU had higher activity on xylan agar (Figure 3.11), whereby large halos were present at 40 °C at pH 6.2 and 8.5, and at 50 °C at pH 8.5. AHU was able to grow and have enzymatic activity a pH 11 from 40-60 °C and was therefore considered thermotolerant and alkaphilic (Rothschild & Mancinelli, 2001). No visible growth was detected at pH 6.2 and 50 °C but there was considerable activity. This could show a large secretion of enzyme into

the media to gain sugars required for growth. Though activity was lower on starch agar, there was still considerable enzymatic action at 70 °C (Table 8.4). There were large amounts of bacterial growth at lower temperatures and near neutral pH. However, enzyme activity was proportionately less; enzymes may not have been secreted, but retained in the cytosol or were cell bound (Golyshina, Golyshin, Timmis, & Ferrer, 2006).

Another strain selected for whole genome sequencing (AHW), was identified as *Brevibacillus agri* via 16S rRNA amplicon sequencing. The genus *Brevibacillus* has been identified as an aerobic, thermophilic, cellulolytic strain which is able to utilise CMC and xylan (Liang *et al.*, 2009; Logan *et al.*, 2002; Maki, Leung, & Qin, 2009). *B. agri* has also been successfully used in microbial fuel cells (Drapcho, Nhuan, & Walker, 2008). Xylan agar yielded only one point of positive enzymatic activity at pH 11 and 50 °C (Figure 3.12). However, the isolate was able to grow at 40-60 °C across the pH range, from 6.2-11. Surprisingly, AHW had no visible activity on starch agar, though it was able to grow between pH 6.2 and 8 across the whole temperature range. Again, such activity must be due to slow acting endo-acting enzymes or utilisation of the basal medium.

Due to their activity at high pH and temperature across the carbon range, 11 isolates that were chosen for genome analysis, were identified as potential *B. licheniformis* strains (Table 3.2). NCBI list 17 *B. licheniformis* genomes that have been sequenced; four complete, two to scaffold level and 11 to contig level (National Center for Biotechnology Information, 2015). Some strains have been reported to have endoglucanases from families GH5 and GH9, and a probable cellulose-1,4-beta-

cellobiosidase (family GH48). *B. licheniformis* have been shown to break down cellulose (mostly in CMC form) for utilisation as a carbon and energy source (Rey *et al.*, 2004). Potential *B. licheniformis* strains (AHY1, AIB, AIE1, AJF1, AJK, AJN1a, AJN1b, AJT and AJVa) had a lot of activity at pH 6.2 and 8, especially at 40-60 °C on CMC agar. An isolated endo-β-1,3-1,4-D-glucanase has been reported to have an optimal temperature of 55 °C, and a wide pH range of 4-10.5 (Lloberas, Querol, & Bernues, 1988). The pH range is reflected in strain B-41361 endoglucanase, though optimal temperature was 65 °C (Bischoff *et al.*, 2006). Of the isolates identified as potential *B. licheniformis*, eight had activity at 70 °C (AHY1, AIB, AIE1, AJG, AJN1a, AJN1b, AJT, and AJVa), of these optimal pH was 6.2 and/or 8. However, five *B. licheniformis* strains were active at pH 10 on CMC agar (AJG, AJK, AJN1a, AJN1b and AJT). These isolates were active at 40 °C on CMC; a single isolate was active at 50 °C (AJN1b) and one at 60 °C (AJN1a). *B. licheniformis* GXN151 has been shown to have a high growth rate at 60 and 70 °C, and create halos on Avicel media (Y. Liu, Zhang, Liu, Zhang, & Ma, 2004). *B. lichenifomis* SVD1 has been shown to contain cellulosomes (extracellular enzyme complexes where enzyme subunits are attached via scaffoldin proteins, through dockerin amino acid sequences), as well as xylanolytic multi-enzyme complexes, though xylan degrading activity was predominant (van Dyk *et al.*, 2009). Interestingly there was proportionately less growth than activity across the carbon utilisation profiles of each isolate, when assayed on CMC agar. Such behaviour could indicate enzyme secretion into the media (Y. S. Kim, Jung, & Pan, 2000; C. C. Lee, Wong, & Robertson, 2001; Linger, Adney, & Darzins, 2010; Stinson & Merrick, 1974) rather than cellulosome or cell-bound enzymes.

B. licheniformis is listed as one of the top natural bacterial xylanases, as documented in the BRENDA database (N. Liu *et al.*, 2011); however *B. pumilus* and *B. subtilis* have greater specific activity (Technische Universitat Braunschweig, 2014). It has been reported that some strains of *B. licheniformis* have an optimum of pH 6-7 in xylan assays, and an optimal temperature of 55 °C (van Dyk *et al.*, 2010). Others retain hydrolytic activity up to pH 11, with optimums of 7 or 8-10 (Damiano *et al.*, 2006). The optimal temperature of isolates on xylan agar appeared to be 50 °C. Of the 11 thermotolerant potential *B. licheniformis* isolates, nine had activity at pH 11 at 50 °C; only AIK did not, but was active at pH 6.2 and 8.5 at this temperature (Figure 3.17). Though AIK was not active at 70 °C, the remaining nine isolates were active at pH 6.2 at this temperature, with AIB, AJN1a, AJN1b and AJT also active at pH 8.5 at 70 °C. This indicates both thermotolerance and alkaline tolerance.

B. licheniformis' amylases have been researched thoroughly and are used industrially for starch degradation, due to their stability at high temperatures (Declerck *et al.*, 2003; Gray *et al.*, 1986; Hmidet *et al.*, 2008). Therefore, it would be expected that results from the starch assays would have a large coverage of temperature. Of the 11 potential thermotolerant *B. licheniformis* strains, five were not active at 70 °C. AIK was the only isolate with no activity on starch agar, however there was growth present at pH 6.2 across the temperature range. There was proportionately more growth than activity for all isolates on starch agar, thus the enzyme could be present in the cytosol rather than secreted (Honda & Kitaoka, 2004; Martoglio & Dobberstein, 1998). All nine active isolates had halos at pH 9.

AIF was identified as potentially a *Bacillus stratosphericus* and was isolated at 40 °C. There was no activity on either CMC or starch at any temperature or pH. However, there was considerable growth on the plates. Activity on xylan was low but consistent across the pH range at 40 °C. Though an increase in temperature caused a decrease in activity, which became none existent at 60 and 70 °C. This was however, not unexpected due to the isolation temperature of this bacterium. The growth of AIF on xylan was considerably higher than the activity, and thus the enzymes used were most probably cell bound or within the cell.

Based on their carbon utilisation profiles four isolates chosen for whole genome sequencing (AHU, AIL, AJF2 and AJH1) were identified (via 16S rRNA amplicon sequencing) as potential *Bacillus subtilis* strains. *B. subtilis* starch and maltose/maltodextrin enzymes have been thoroughly characterised, especially alpha-amylases (Asgher, Asad, Rahman, & Legge, 2007; Bano, Ul Qader, Aman, Syed, & Azhar, 2011; Mukherjee, Borah, & Rai, 2009). However, a number of other enzymes, such as beta-glucosidases (LicH), are able to depolymerise β-1,3-1,4-glucan and other glucans (Schönert *et al.*, 2006; Tobisch, Glaser, Krüger, & Hecker, 1997), thereby efficiently utilising plant biomass for glucose monomer release (Crittenden *et al.*, 2002; Pieper *et al.*, 2008). At 40 and 50 °C, all four isolates were active on starch agar from pH 6.2-9 (Figure 8.3). At 60 °C, activity dropped and was restricted to pH 8, bar isolate AJF2 where activity was seen at pH 9 (Table 8.5). Only AJH1 had no activity at 70 °C, AHU, AIL and AJF2 were still active at pH 8 at this temperature (Table 8.5). High temperature activity at pH 8 is not unusual and thermostability is desired industrially for enzymes (Asgher *et al.*, 2007).

Thermostable alkaline cellulases have been isolated from *B. subtilis* MBG874 (van Dijl & Hecker, 2013). All potential *B. subtilis* isolates chosen for whole genome sequencing had activity across pH 6.2-10 at 40 °C on CMC agar (Table 8.5). Isolates AJF2 and AJH1 were both active up to pH 10 at 50 °C, whereas AHU and AIL were only active to this pH at 40 °C. Activity was highest at 40 °C for AHU and AJH1 across the carbon sources, whereas optimal temperature for AIL and AJF2 was at 50 °C. All isolates were still active on CMC agar at 70 °C, though only at pH 6.2 and 8. Withstanding this pH and temperature range with only CMC as a carbon source, such isolates could be of use in the fermentation industry (Schallmey *et al.*, 2004). There were large amounts of bacterial growth at the lower temperatures and near neutral pH. However, enzyme activity was proportionately less, whereby enzymes may not be secreted but instead retained in the cells cytosol or displayed on the cell surface (Golyshina *et al.*, 2006; Y. S. Kim *et al.*, 2000).

Alkaline endo-xylanases have been isolated from *B. subtilis* (Schallmey *et al.*, 2004) which is known for xylan utilisation (Roncero, 1983). Xylanases generally have an optimal temperature near 50 °C and pH of 4-6 (Bernier, Desrochers, Jurasek, & Paice, 1983; Biocatalysts Ltd., 2015a, 2015b; Novozymes Bioenergy, 2010b). However, xylanases have been found with optimal pH up to 8 at this temperature (Kamble & Jadhav, 2012) or pH 4-6 at 60 °C (G. Guo *et al.*, 2012). Bar AIL, all potential *B. subtilis* isolates chosen for whole genome sequencing were active up to pH 11 at 40 and 50 °C on xylan agar (Figure 3.11, Figure 3.19, Figure 3.21 and Table 8.5). However, AIL was active from pH 6.2-11 at 60 °C. All four isolates were active at pH 6.2 at 70 °C; AHU and AJH1 were also active at pH 8.5 at this temperature. Such alkaline tolerant enzymes,

with a broad temperature range, would be useful in food processing (Schallmey *et al.*, 2004). With activity halos proportionally larger than growth, xylanases may have been secreted into the agar (Ruller *et al.*, 2006).

Weak acid hydrolyisis of sugars using HCl, H_2SO_4, HF and CH_3COOH, is used in industry as a pre-treatment for lignocellulosic biomass (David, Fornasier, Greindl-Fallon, & Vanlautem, 1985; Larsson *et al.*, 1999; Lavarack, Griffin, & Rodman, 2002). These acids are able to destroy heterocyclic ether bonds between hemicellulose and cellulose molecules, by releasing protons. Some of the bacterial isolates may not have specific enzymes to degrade the carbon sources but in fact use acid hydrolysis to liberate glucose, xylose or fructose (Aguilar, Ramirez, Garrote, & Vazquez, 2002). This could also influence survival on alkaline media, as the production of acetic acid from hydrolysis of acetyl groups, which can enter the cells, lowers the internal pH (Aguilar *et al.*, 2002; Larsson *et al.*, 1999).

Mixed cultures in the bacterial library may communally break down substrates at a lower pH, as most isolates that were active on pH media were not pure cultures. In the mesophilic collection 17.3% of isolates were mixed, whereas only a single isolate was considered mixed in the thermotolerant library (Table 8.2). pH has been shown to alter composition of mixed communities allowing those that can tolerate higher or lower pH to proliferate (Bradshaw & Marsh, 1998), though diversity can also increase with lower pH (Fang & Liu, 2002). It has also been reported that, during a drop in pH, energy was diverted to none growth functions, such as carbohydrate utilisation, rather than protein synthesis (Strobel & Russell, 1986). This would therefore be detrimental to protein recovery, but useful for platform chemical production. The use of

thermophiles can raise energy costs, due to the maintenance temperature for optimal activity. However, hydrothermal pretreatments of biomass mean that the fermentation feedstock is already hot, plus sanitation wastes could be utilised due to their requirement of 1 hour heat treatment at 70 °C (Kongjan *et al.*, 2010). Also, efficient production of hydrogen from *Miscanthus* using moderately alkaliphilic and extreme thermophilic bacteria *Caldicellulosiruptor saccharolyticus* and *Thermotoga neapolitana* has been reported. By pre-treating material with mild alkaline, fermentation inhibitors are avoided and high temperatures ensure isolates utilise all available sugars, and hydrogen is produced (Truus de Vrije *et al.*, 2009). Therefore, by selecting isolates able to utilise xylan and cellulose at high pH and temperature for protein analysis (section 4.2.8) and whole genome sequencing (section 5.3), there may be greater opportunity to discover an isolate for biorefining application.

3.3.1 Conclusions

A wide range of bacteria were isolated from degrading *Miscanthus* chip. Many were identifiable using amplicons, produced by primers specific for the 16S hypervariable regions of small subunit rRNA. Though 16S rRNA amplification and sequencing was relatively quick and cheap compared to other methods, it would have been useful to try other methods for identification. Some bacteria were not identifiable, as many sequences were contaminated or unusable, which may have been due to the presence of mixed cultures. Much PCR optimisation took place for 16S amplification before products could be sequenced, so the process was not as quick as expected. Morphological assessment aided bacterial identification, especially when Gram staining did not appear correct. Isolates differed considerably under selection at higher

temperature, resulting in more thermotolerant species such as *Bacilli*. The more extreme the isolation conditions were the fewer species were found.

Fosmid libraries were a consideration, but would not have been as useful for isolating bacteria that could be used in biorefining. However, more species would undoubtedly have been identified via sequencing, as varied culture conditions would not have been necessary due to fosmids being expressed in *Escherichia coli.* Also, enzymes may have been easier to screen for and isolate when using *E. coli* with optimum conditions, as overgrowth of neighbouring colonies would not have occurred.

Bacterial isolates that were cultured from *Miscanthus* chip at room temperature had a range of enzymes hypothetically capable of degrading plant cell walls. Many isolates were able to withstand a range of pH, mainly at the higher more alkaline end of the spectrum. Due to their ability to withstand extremes of pH when degrading xylan 23 isolates were chosen to be studied more in depth. A number of isolates were still able to grow and active at extremes of pH and temperature combined; such characteristics would be useful in current saccharification and fermentative technologies.

Chapter 4 Enzyme Isolation and Expression

4.1 Introduction

The aim was to isolate genes encoding alpha-amylase and xylan deacetylase enzymes, express them and characterise the activity. A cytoplasmic alpha-amylase identified in *Bacillus coagulans* ABQ genomic DNA was selected, amplified, cloned as a 6x-His fusion protein into *E. coli*, and induced for expression. A further objective was to demonstrate secretion of cellulase and xylanase enzymes into media by a number of bacterial isolates, which performed well in the carbon utilisation screens (section 3.2.5). This aimed to prove that little growth of bacterial colonies, which produced a large halo in the carbon utilisation profiles (section 3.2.5), was due to secretion of enzymes.

Isolation and characterisation of bacteria (Chapter 3) lead to the isolation of an alpha-amylase gene. ABQ was identified as *Bacillus coagulans* via partial 16S rRNA amplicon sequencing (section 3.2.1). Published whole genome sequences of *B. coagulans* strains were used to design primers for amylases, xylanases and cellulases that would be of use for plant biomass biorefining. *B. coagulans* has been described as a Gram positive rod (0.5-2.5 x 1.2-10 μm in size). Straight rods can be found in pairs or chains with either rounded or square ends (Holt *et al.*, 2000; Rhee *et al.*, 2011). Cells are motile via peritichous flagella and endospores are formed, which could withstand extreme conditions. *B. coagulans* is a facultative anaerobe, which is catalase positive and can tolerate a range of temperatures (Holt *et al.*, 2000). It is classed as a thermophilic, lactic acid producing bacteria (Batra, Singh, Banerjee, Patnaik, & Sobti, 2002), with an optimal growth temperature of 55 °C and pH of 5.5 (Payot, Chemaly, &

Fick, 1999; Rhee *et al.*, 2011). Strains identified in section 3.2.1 shared similar characteristics with this description. The three *B. coagulans* strains used here were isolated at 40 °C and pH 5.5 from soil around a *Miscanthus giganteus* plot. The enzymes obtained from these bacteria were expected to be relatively thermotolerant and acidophilic due to the isolation conditions. Available on GenBank were three complete genomes and nine assemblies at scaffold or contig level of *B. coagulans*. Primers based on genes encoding enzymes (Table 2.7) were designed based on two of the published genomes (2-6 (Su *et al.*, 2011), which NCBI refers to as the "representative", and 36D1 (Rhee *et al.*, 2011)). Both genomes were fully assembled and annotated (NCBI, 2014), and used to design primers for candidate genes, which were amplified from the isolated strains. However, the 36D1 genome had xylanases that were not present in the 2-6 genome, which could have been gained from horizontal transfer (Amábile-Cuevas & Chicurel, 1993; M Syvanen, 1994). This may have occurred via transfer of a cassette of enzymes into the genome, as the genes encoding the enzymes were positioned closely together (R. M. Hall, 2002; Recchia & Hall, 1997). Primers were designed to identify whether the three *B. coagulans* strains isolated from *Miscanthus* plots had genes encoding amylases, xylanases and cellulases. Xylanases could be isolated and potentially used to metabolically engineer another organism for the production of xylose, xylitol etc. However, the isolated strains of *B. coagulans* were found to have genes encoding a starch degrading enzyme (alpha-amylase), which was purified and characterized in this chapter.

Genomic DNA was isolated from the three *B. coagulans* strains for amplification of genes, followed by cloning into the pGEM-T vector and transformed into *E. coli* for

blue/white screening. Positive transformations were cloned into pET30a vector and transformed into *E. coli* BL21Star for expression of the alpha-amylase protein. IPTG induction was used for expression of the recombinant fusion protein followed by purification using Ni-NTA. The recovered protein was desalted for protein quantification and characterisation.

4.2 Results

Genomic DNA was isolated from *B. coagulans* ABP, ABQ and ABR using the DNeasy Blood and Tissue Kit (Qiagen, Redwood City, 69581; section 2.2.2.2). Enzyme gene primer combinations (Table 2.7) were used to screen for enzymes (xylanase cassette, alpha-amylase and endo-glucanase) within each bacterial strain by PCR amplification. Most primer pairs did not amplify a product. The most likely cause is that they were not actually present in the genome, as they were based on what was most likely a gene cassette in the 36D1 genome (Amábile-Cuevas & Chicurel, 1993; R. M. Hall, 2002; M Syvanen, 1994). Alternatively, primer binding sites may have differed compared to 36D1 and 2-6 strains. However, surprisingly one strain (*B. coagulans* ABQ) appeared to have an endoglucanase present as well as the polysaccharide/xylan deacetylase, which was present in all three strains. A restriction digest of pGEM-T vector containing the alpha-amylase gene occurred using *XbaI* and *XhoI* enzymes (section 2.2.5). Two strains of *B. coagulans* (ABP and ABQ) were chosen due to complete alpha-amylase sequence information gained from plasmid sequencing with T7 promoter and SP6 primers. *B. coagulans* ABR had single nucleotide polymorphisms (SNPs) (Brookes, 1999) or whole sections missing from the PCR amplicon sequence. The alpha-amylase gene was cloned into pET30a and positive transformations were selected for colony PCR with the T7

terminator and promoter to ensure the insert and plasmid were both present. The bacteria were then lysed by sonication and lysozyme action for solubility tests (section 2.2.6.1.2) and the recombinant fusion protein purified for further characterisation.

4.2.1 SDS-PAGE gel for Pilot Expression solubility test

Alpha-amylase was calculated to be 56 kDa (Swiss Institute of Bioinformatics, 2015) in size which was confirmed SDS-PAGE (Figure 4.1). The IPTG induced samples were indicated by thick, dark bands which showed that there was a larger amount of protein of this size being expressed. Though *E. coli* also has an alpha amylase, it has a different sequence and is larger than the *B. coagulans* gene (Freundlieb & Boos, 1986) and should not be expressed as much as the induced gene. A larger product (99.78 kDa, Figure 4.1) was also expressed to the same degree as the alpha-amylase. This could indicate aggregation (Bondos & Bicknell, 2003) or *E. coli* expressed part of the vector as well as the gene.

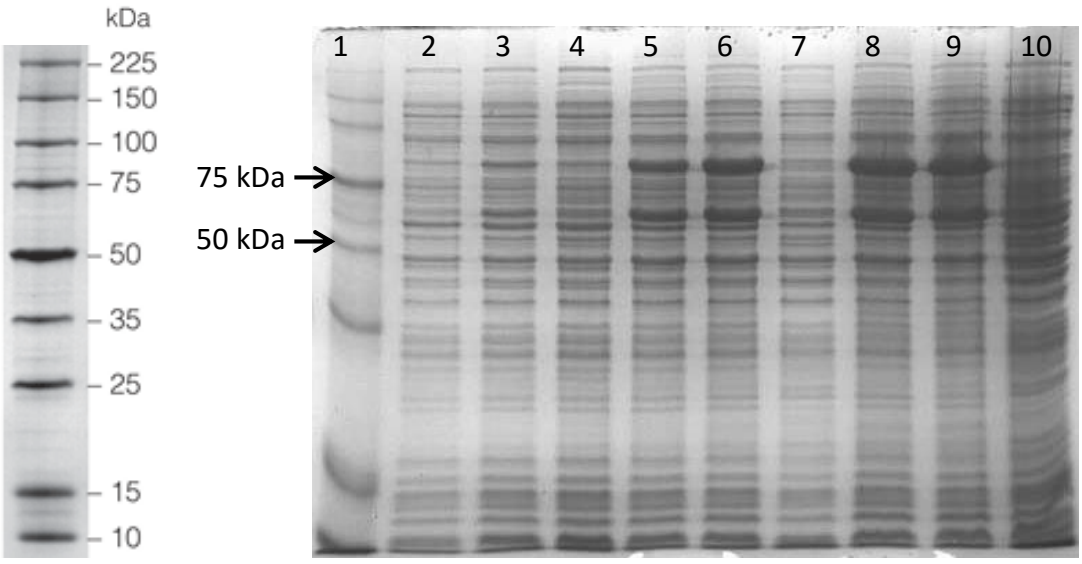


Figure 4.1 Coomassie stained 12% SDS-PAGE.

Showing alpha-amylase expressed in *E. coli* BL21 Star with pET30a expression vector from *B. coagulans* strain ABQ. From the sequence the alpha-amylase was calculated to be 56 kDa and can be seen as the thicker band above 50 kDa. *1- Broad range*

molecular marker (Promega, V8491, image from promega.co.uk); 2 - time zero alpha-amylase expressed in E. coli BL21 Star with pET30a expression vector from B. coagulans strain ABQ, induced with IPTG; 3 - 1.5 hrs induction; 4 – 1.5 hrs non-induced proteins; 5 - 3 hrs induction; 6 - 4.5 hrs induction; 7 – 4.5 hrs non-induced proteins; 8 – 6hrs induction; 9 – 8hrs induction; 10 – 8 hrs non-induced protein.

4.2.2 Confirmation of His-tagged proteins using Western blot

SDS-PAGE was utilised to assess loading needed for the Western blot transfer and probing with the anti-His C-terminal antibody (section 2.2.6.1.4). The antibody probed transfer membrane showed two artefacts that were present (below 50 kDa). A small amount of artefact was transferred; therefore the amino acid sequence must have contained six histidine residues. The Western blot membrane (Figure 4.2) shows that both large bands were transferred and contain the alpha-amylase gene as the antibody was specific to the C-terminus Histidine tag on the pET30a vector (Anti-His(C-term)-AP Antibody from Invitogen, Cat. No. R932-25). Ponceau S (Sigma-Aldrich, P3504-50G) was used as a total protein pre-antibody stain for normalisation.

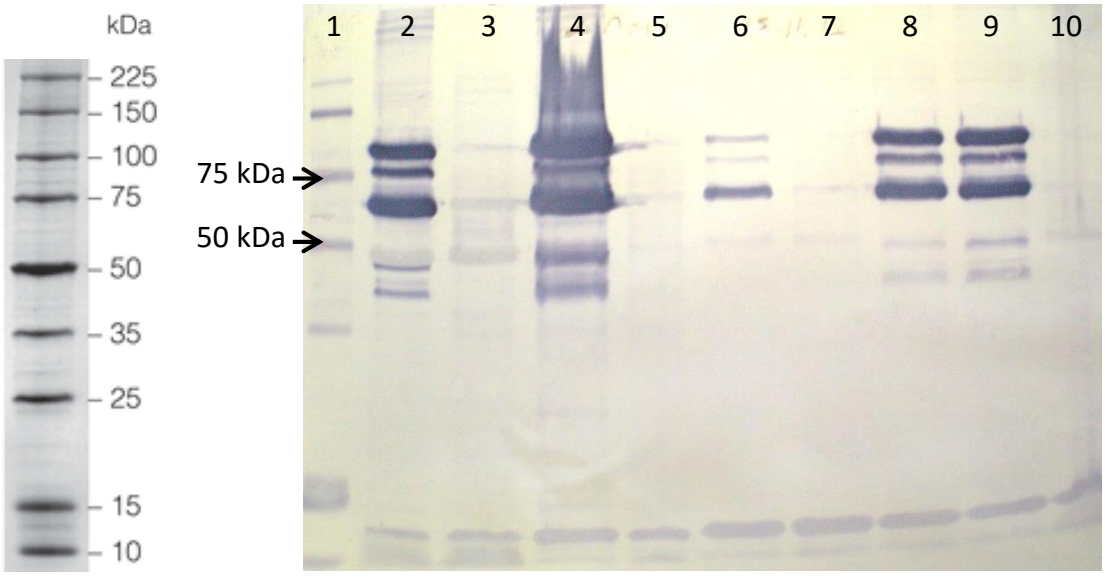


Figure 4.2 Western Blot membrane.

IPTG induced *E. coli* BL21 Star for expression of alpha-amylase from *B. coagulans* strain ABQ, probed using Anti-His(C-term)-AP Antibody from Invitogen (Cat. No. R932-25). *1 – Broad range molecular marker (Promega, V8491, image from promega.co.uk), 50*

kDa weighted at 0.3 μg/μL; 2 – 8 hrs IPTG induction, concentrated pellet sample; 3 - 8 hrs non-induced, concentrated pellet sample; 4 - 8 hrs IPTG induction, concentrated supernatant sample; 5 - 8 hrs non-induced, concentrated supernatant sample; 6 – 1.5 hrs IPTG induction, supernatant sample; 7 - 1.5 hrs non-induced, supernatant sample; 8 - 4.5 hrs IPTG induction, supernatant sample; 9 - 6 hrs IPTG induction, supernatant sample; 10 - 6 hrs non-induced, supernatant sample.

4.2.3 Ni-NTA Purification SDS-PAGE gels

Figure 4.3 shows most *E. coli* proteins were removed from the sample using nickel-nitrilotriacetic acid (Ni-NTA) purification (section 2.2.6.2) though consecutive elutions were not totally pure. Further purification would have been required to obtain a pure protein product. The alpha-amylase enzyme could be seen clearly in the elutions.

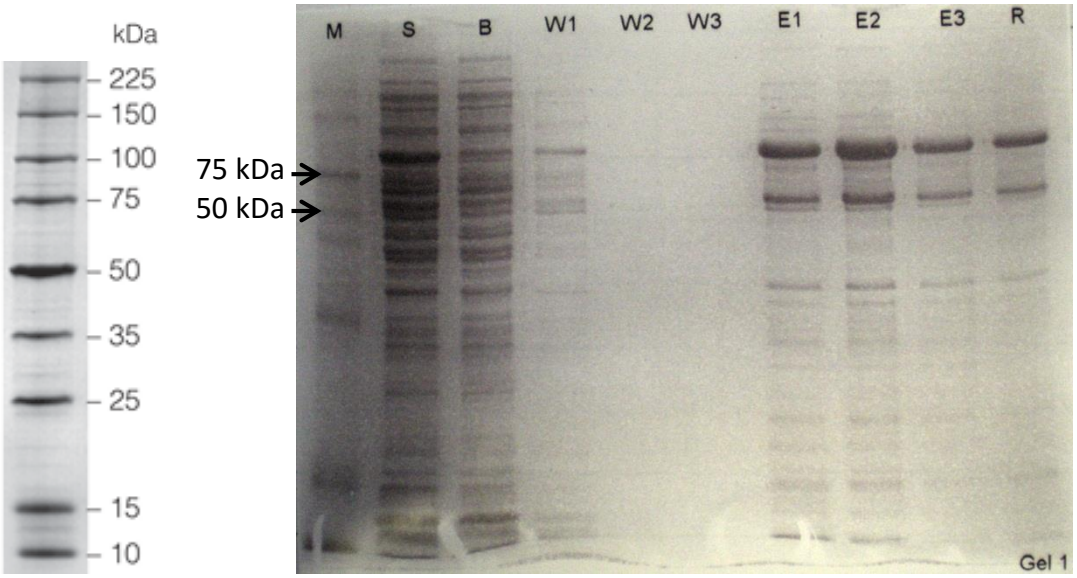


Figure 4.3 Coomassie stained 12% SDS-PAGE showing Ni-NTA purification.

C-terminal His-tagged alpha-amylase protein from *B. coagulans* ABQ, which had been expressed in *E. coli* BL21 Star using the pET30a expression system. *M – molecular marker (Promega, V8491, image from promega.co.uk), B – bound to resin, W1 – wash 1, W2 – wash 2, W3 – wash 3, E1 – elution 1, E2 – elution 2, E3 – elution 3, R – resin sample.*

4.2.4 Large scale production of recombinant alpha-amylase from *B. coagulans* ABQ

Expression of alpha-amylase from *B. coagulans* ABQ using pET30a expression vector in *E. coli* BL21 Star was scaled to 1 litre cultures to obtain a larger amount of protein for

characterisation. Growth had been slow and had not reached OD_{600} 0.5-0.8 with enough time for 4 hrs of expression, therefore flasks were left shaking overnight at room temperature (final OD_{600} of 3.36). A new inoculum was made from this overnight culture for the experiment to start the following day. After 2 hours of incubation an OD_{600} of 0.08 was recorded. This was too low to begin induction with IPTG, so 1% of overnight culture was added to boost optical density (OD). However, due to continual slow growth, 100 mL of the overnight cultures had to be used to boost OD of the new 500 mL cultures. After 4 hrs of IPTG induction (Table 4.1), the cultures were centrifuged to gain the *E. coli* pellet containing the alpha-amylase protein. Expression was "leaky" as the expressed fusion protein was found in both the supernatant and pellet (Figure 4.2). Alternatively, this could have been the soluble fraction of the bacterial pellet. Pellets were then weighed, sonicated and Ni-NTA purified.

Table 4.1 Optical densities, at 600 nm, of the IPTG induced scaled up pilot expression.

Alpha-amylase from *B. coagulans* ABQ using pET30a and *E. coli* BL21 Star for a 1 L culture. A non-induced sample was not taken through the time points as it was known from previous expression investigations that only 4 hours of induction with IPTG were required for a large amount of protein to be produced. ** - Rotation had been switched off before the time point 3, thus little growth.*

Sample	Time	Time Point	OD_{600}
Flask 1	-		0.96
Flask 2	-		0.92
Flask 1	1.30	1	1.26
Flask 2	1.30		1.26
Flask 1	2.30	2	1.36
Flask 2	2.30		1.41
Flask 1	4.00	3	1.36*
Flask 2	4.00		1.42*
Flask 1	5.30	4	1.48
Flask 2	5.30		1.52

4.2.5 Zymography gel next to a coomassie stained gel

A modified native-PAGE protocol was used to view activity of the purified alpha amylase protein, whereby SDS-PAGE without SDS and βME (beta-mercaptoethanol) was run at a cooler temperature by placing an ice pack in the tank. The gel was split (first section for coomassie, second section for zymography) before treatment with wash buffers to renature the proteins and stained with iodine (section 2.2.6.6). Figure 4.4 A) shows that the alpha-amylase gene was present. Though it had been thoroughly boiled before loading the negative control (bacterial alpha-amylase, Fisher) was still active on the zymography gel. Two bands of activity (arrows) could be seen on the zymography gel (Figure 4.4 B)) indicating that both bands seen previously (Figure 4.3) were both active.

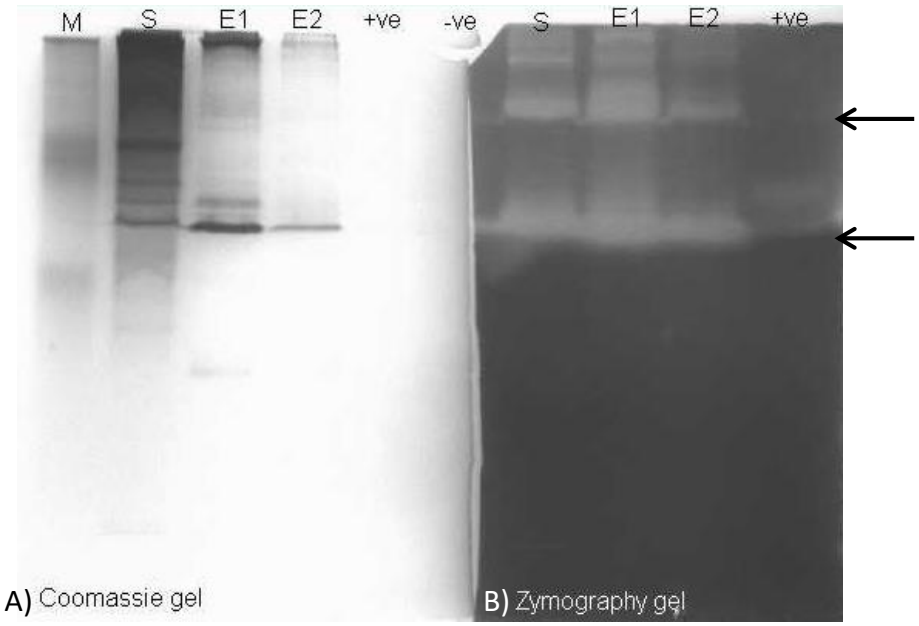


Figure 4.4 Native PAGE and Zymography gel (12% gel containing 1% starch).

Showing activity of alpha-amylase from *B. coagulans* ABQ expressed with pET30a in *E. coli* BL21 Star. A cleared halo appeared where the amylase was active on gel B) shown by arrows. *M – broad range protein marker (BioRad); S – supernatant from sonicated pellet; E1 – Ni-NTA purified alpha-amylase elution 1; E2 – Ni-NTA purified alpha-*

amylase elution 2; +ve – bacterial alpha-amylase at 100 mg/mL (Fisher); -ve – boiled alpha-amylase at 100 mg/mL (Fisher).

4.2.6 Starch Agar assay

A simple starch agar assay took place to support the zymography gel results (Figure 4.4). Figure 4.5 shows that the first and second elutions had larger, clearer halos and were more intense as more alpha amylase was present. Elution 3 had a smaller halo as less protein was present. Figure 4.5 A) with 50 μL loading had more intense, clearer halos, as a higher proportion of enzyme was loaded. However, both plates show concentrated purified alpha-amylase in comparison to the positive controls (1-6), of which only "1" had a slight halo (100 mg/mL of bacterial alpha-amylase). The T5 induced (T5I) and non-induced (T5NI) lysed cell samples supported the elution's halos, as T5I had activity whereas T5NI did not. The non-induced sample did not have enough alpha-amylase present to break down the starch, whereas the induced sample had a small halo signifying starch degradation.

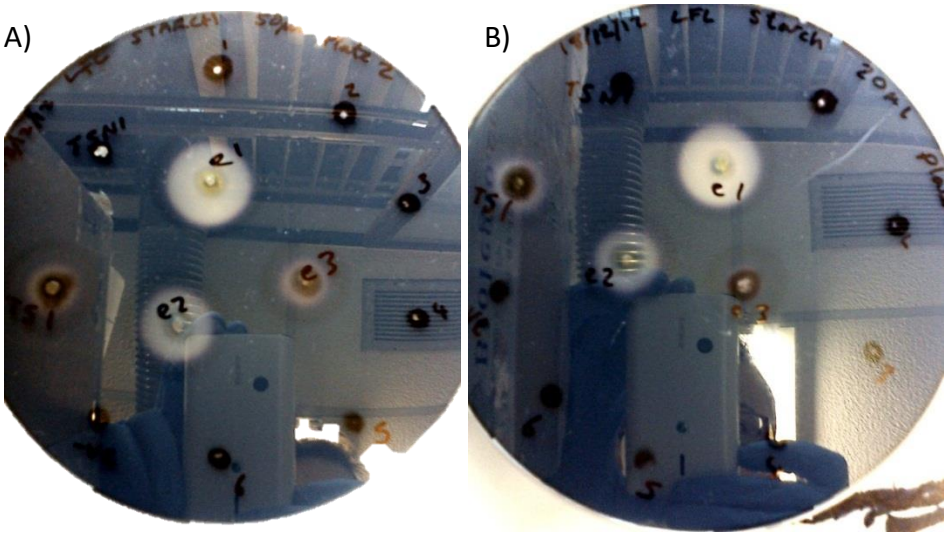


Figure 4.5 Simple starch assay on 1% starch agar

Purified *B. coagulans* ABQ alpha-amylase for starch depolymerisation activity. *A) – 50 µL loading of each sample; B) – 20 µL loading of each sample; 1 - 100 mg/mL of bacterial alpha-amylase (Fisher); 2 – 10 mg/mL alpha-amylase (Fisher); 3 – 1 mg/mL alpha-amylase (Fisher); 4 – 0.1 mg/mL alpha-amylase (Fisher); 5 – 0.01 mg/mL alpha-amylase (Fisher); 6 – 0.001 mg/mL alpha-amylase (Fisher); -ve – boiled alpha-amylase at 100 mg/mL (Fisher); T5 I – lysed E. coli BL21 Star containing ABQ alpha-amylase, induced to time point 5 with IPTG; T5 NI – lysed E. coli BL21 Star containing alpha-amylase, non-induced with IPTG at time point 5 of growth; e1 – elution 1 of Ni-NTA purified ABQ alpha-amylase; e2 – elution 2 of Ni-NTA purified ABQ alpha-amylase; e3 – elution 3 of Ni-NTA purified ABQ alpha-amylase.*

4.2.7 Alpha-Amylase Characterisation

The alpha-amylase isolated from *B. coagulans* in this chapter was further characterised with preliminary tests using a number of assays to test the pH range, temperature range, thermostability and effect of cofactor. Firstly, the enzyme was tested using a Phadebas Amylase Test to assess which conditions were required for activity, before a larger amount of protein was acquired for further characterisation assays.

4.2.7.1 Phadebas amylase test

Phadebas Amylase Test was utilised to measure the activity of the selected enzyme, over a time course, with an alpha-amylase specific assay with set conditions. This ensured that the enzyme was active, and could therefore be subject to further assays (sections 4.2.7.3 to 4.2.7.6). The assay occurred at 50 °C after assaying at the recommended temperature of 37 °C resulted in very little activity. The pH was neutral and distilled water was to allow for Ca^{2+} ions to perform as a cofactor for the enzyme. Figure 4.6 shows the mean (n=4) absorbance readings at 600 nm for the assay. After completion of the assay, the supplied standard curve (Appendix 8.2.1) was used to work out amylase activity in a variety of units. By assaying a range of eluted alpha-amyalse preparations, a suitable OD was devised for comparison to the positive control at 1 mg/mL (Table 4.2 compared to Table 4.3) throughout further assays (section 4.2.7.2). However, a new batch of protein was required for further assays, therefore the results shown in Table 4.2 did not apply to any of the subsequent assays (bar positive control). ABQ1 and ABQ20 were grown and induced for the same amount of time (4 hours), therefore these had similar results. However, ABQ1o/n had a larger

biomass (as it had been cultured overnight) and therefore contained a lot more protein.

Table 4.2 Phadebas optical density (OD_{600}) readings.

From 1 hour reaction incubation converted to various units using the standard curve supplied with Phadebas Amylase Test (Magle). *ABQ1 – supernatant from 500 mL induced pellet, ABQ1o/n – supernatant from 500 mL induced pellet overnight incubation, ABQ50 – Ni-NTA purified supernatant from 50 mL induced pellet, neg – negative control which had gone through Ni-NTA purification, pos – positive control of alpha-amylase at 1 mg/mL (Fisher) in Ni-NTA elution buffer.*

Name	**OD_{600}**	**Mean /4**	**Standard Curve (U/l)**	**mkat/l**	**Somogyi units/100mL**
ABQ1	0.4975	0.124375	84	1.5	45.36
ABQ1o/n	0.9525	0.238125	190	3	102.6
ABQ20	0.38	0.095	80	1.4	43.2
Neg	0.0025	0.000625	0	0	0
Pos	0.7975	0.199375	180	2.8	97.2

Figure 4.6 shows that as ABQ1o/n had a larger amount of protein present, there was a larger amount of activity over the time course, which plateaued after 4 hours of incubation. As ABQ1 and ABQ20 contained less protein, they were slower to react and were still not at saturation after 4 hours. Surprisingly, the positive control was also relatively slow at utilising all of the substrate and had not plateaued after 4 hours; however, conditions were not optimised for this enzyme. The Phadebas assay showed that the alpha-amylase was active, but further assays would have to take place for up to 4 hours for activity to be recorded.

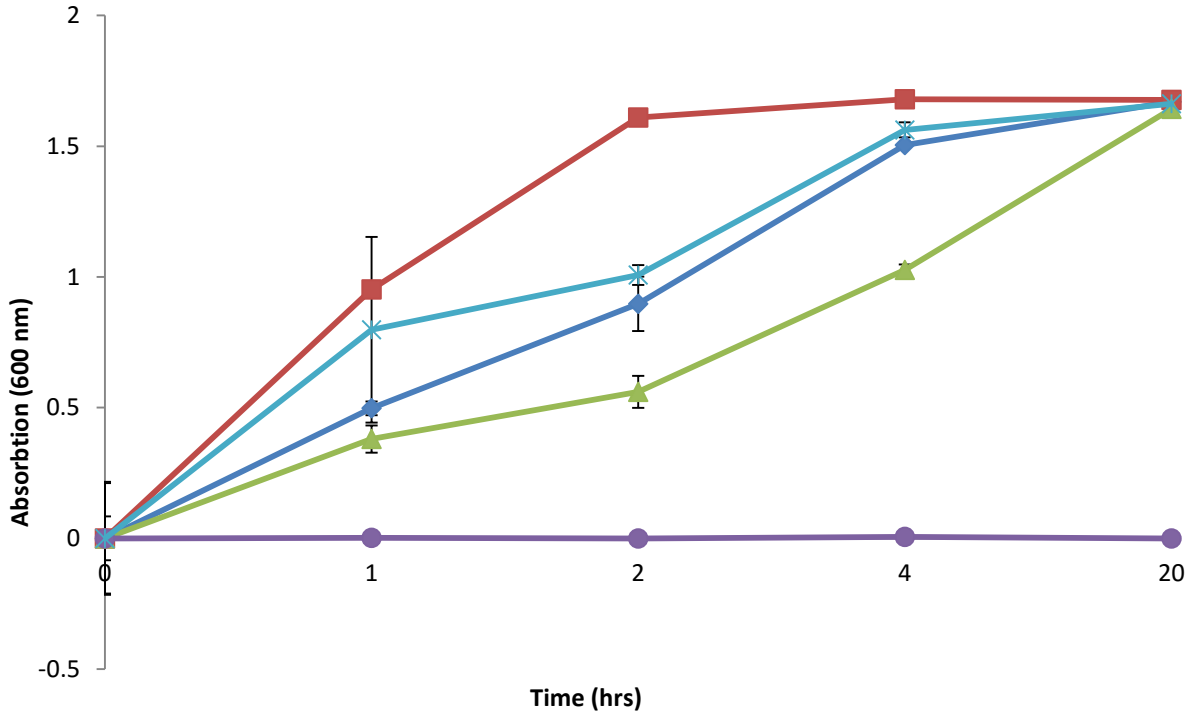


Figure 4.6 Alpha-amylase Phadebas assay.

Amylase had been isolated from *B. coagulans* and expressed using *E. coli* BL21 Star and pET30a expression system had been Ni-NTA purified. Activity was measured using Phadebas amylase test over a time period with set conditions of 50 °C and set pH. Error bars show standard deviation, n=4. *ABQ1 (♦) – 500 mL induced culture; ABQ1o/n (■) – non-induced 500 mL overnight culture; ABQ20 (▲) – 20 mL induced culture; neg (●) – negative control (containing no enzyme); pos (✳) – positive control: bacterial alpha-amylase (Fisher) at 1 mg/mL.*

4.2.7.2 Protein quantification for new alpha-amylase batch

Purified protein samples were quantified using BioRad Protein Assay (section 2.2.6.5). Concentration of the positive control for further amylase characterisation was therefore set at 1 mg/mL. The positive control was not set at 2-3 mg/mL due to the additional proteins still found in the purified amylase (Figure 4.3). However, the OD of 0.735 (Table 4.3) was close to that of the positive control (Table 4.2) and therefore may have been comparable throughout the following assays.

Table 4.3 Purified protein concentrations from optical density readings using a standard curve.
Samples were purified with Ni-NTA, elutions assayed, and calculations based on a BSA standard curve (**Figure 4.7**).

Sample	OD_{600}	Reading	µg/µL	µg/mL	mg/mL
Elution 1	0.735	27	2.7	2700	2.7
Elution 2	0.1	4.1	0.41	410	0.41

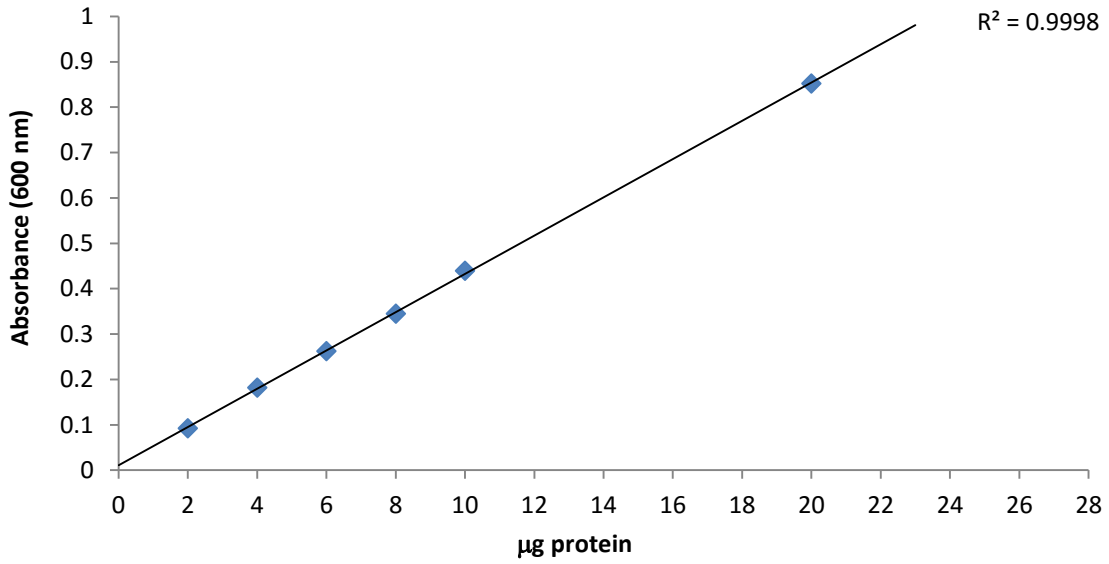


Figure 4.7 Standard curve of absorbance for BioRad Protein Assay using BSA, measured at 600 nm.

4.2.7.3 pH Range

The pH range of *B. coagulans* alpha-amylase was assessed (Figure 4.8) using a range of buffers containing 1% starch at different pHs (section 2.2.6.12, Dawson 2002). The assays were conducted at 50 °C as this appeared to be the favoured temperature of the alpha-amylase, as established by the zymography (section 2.2.6.6 and 4.2.5) and the Phadebas amylase test (section 2.2.6.8). The test also took place across a time scale, to establish the incubation time required for further characterisation assays. Figure 4.8 shows alpha-amylase activity after 30 min, 2 hours, and 4 hours of incubation in each pH buffer (with 1% starch). The lower the optical density the greater activity was present, as more starch was broken down, so there were less long

chain sugars for the iodine to bind to. Activity was most prevalent at two pH values (6.02 and 8.78) with the largest peak at pH 8.78, though these were only obvious after 2 hours of incubation at 50 °C.

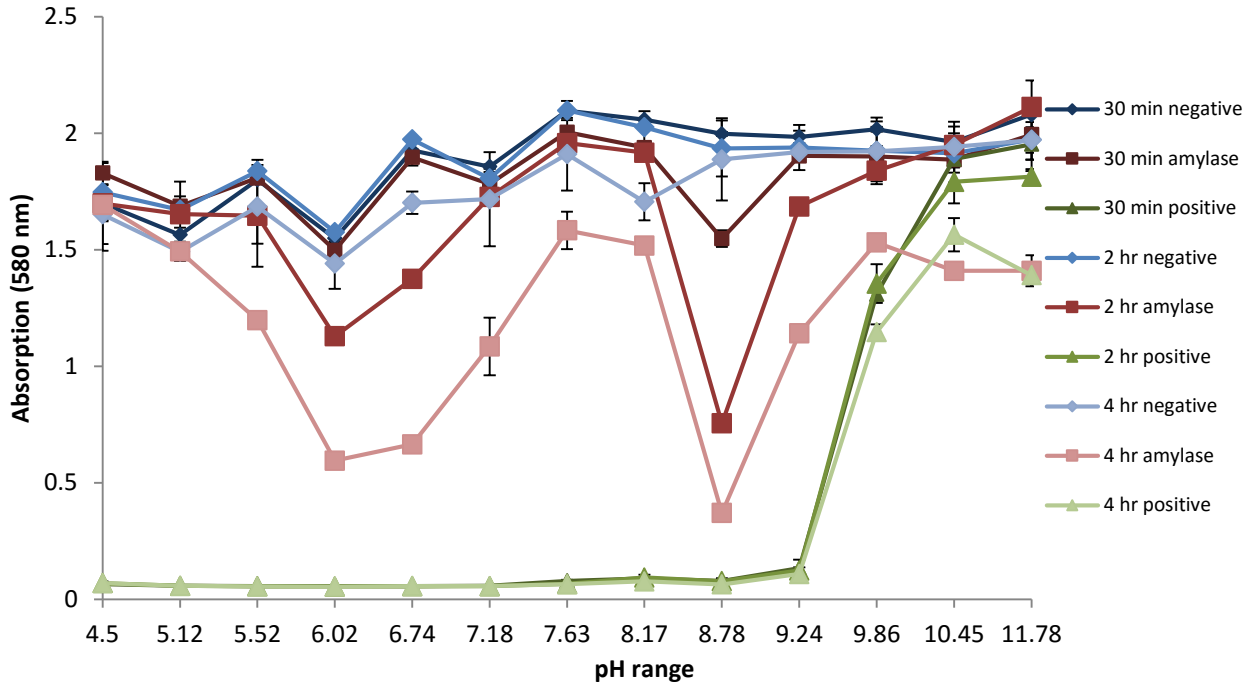


Figure 4.8 pH range of alpha-amylase.

Alpha-amylase isolated from *B. coagulans* and expressed using *E. coli* BL21 Star and pET30a expression system had been Ni-NTA purified, then desalted so that the enzyme could be characterised and here the pH optimum could be found. Activity was assessed via optical density at 580 nm after 30 min, 2hrs and 4 hrs incubation at 50 °C in each pH buffer (containing 1% starch). Error bars show the standard deviation, n=4 for alpha-amylase and n=2 for negative control and positive control. The lower the absorbance the more activity was present as the iodine could only bind to remaining long chain sugars. *Negative control (♦); B. coagulans alpha-amylase (■); positive control (1 mg/mL Fisher bacterial alpha-amylase)(▲).*

4.2.7.4 Temperature Range

The temperature optimum of the alpha-amylase was assessed (section 2.2.6.13) at the optimal pH (8.78) at six temperatures: room temperature, 30 °C, 37 °C (Figure 4.9), 60 °C, 80 °C and 90 °C (Figure 4.10). There was negligible activity at room temperature, which increased with temperature. However, activity declined at higher temperatures

(60-90 °C), bar 7 hours incubation at 60 °C. The lower the optical density reading the greater activity was present, as there were less long chain sugars for iodine to bind to. The negative control fluctuated across the temperature range, whereas the positive control was consistent, especially at lower temperatures.

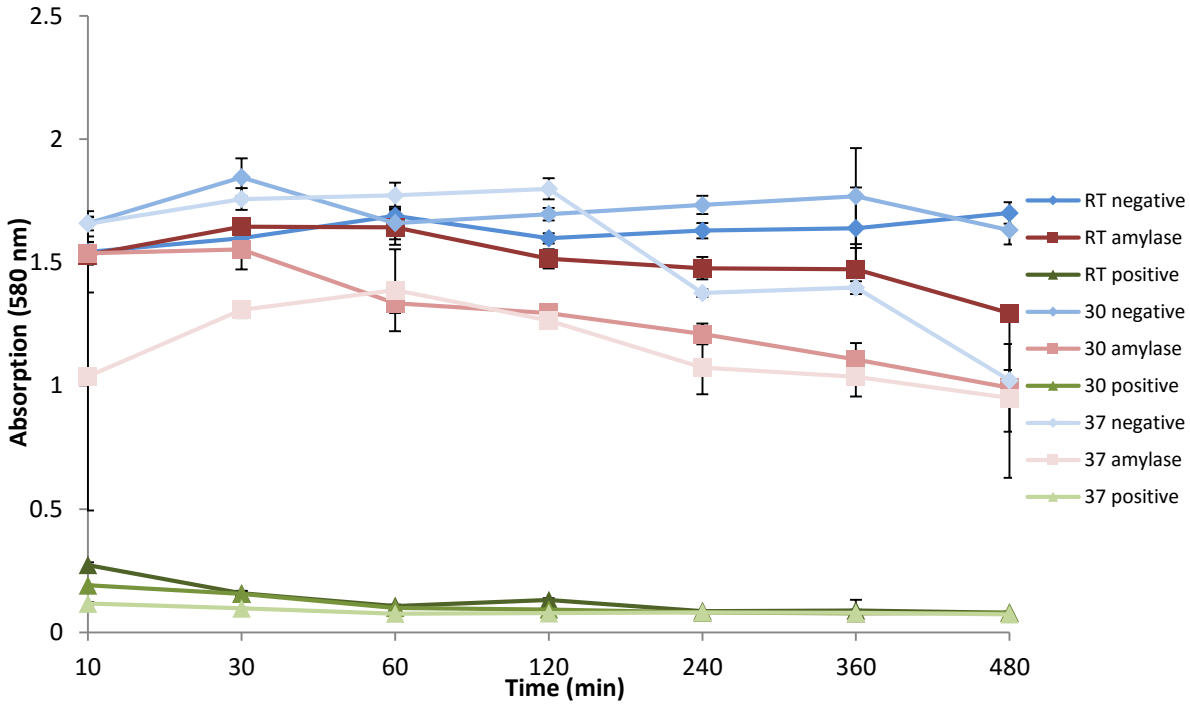


Figure 4.9 Temperature range of alpha-amylase (room temperature to 37 °C).

Alpha-amylase isolated from *B. coagulans* and expressed using *E. coli* BL21 Star and pET30a expression system had been Ni-NTA purified, then desalted so that the enzyme could be characterised and the temperature optimum could be found. Activity was assessed via optical density at 580 nm after 7 hours incubation at room temperature (RT), 30 °C and 37 °C in pH 8.78 buffer (containing 1% starch). Error bars show the standard deviation, n=4 for alpha-amylase and n=2 for negative control and positive control. The lower the absorbance the more activity was present as the iodine could only bind to remaining long chain sugars. *Negative control (♦); B. coagulans alpha-amylase (■); positive control (1 mg/mL Fisher bacterial alpha-amylase)(▲).*

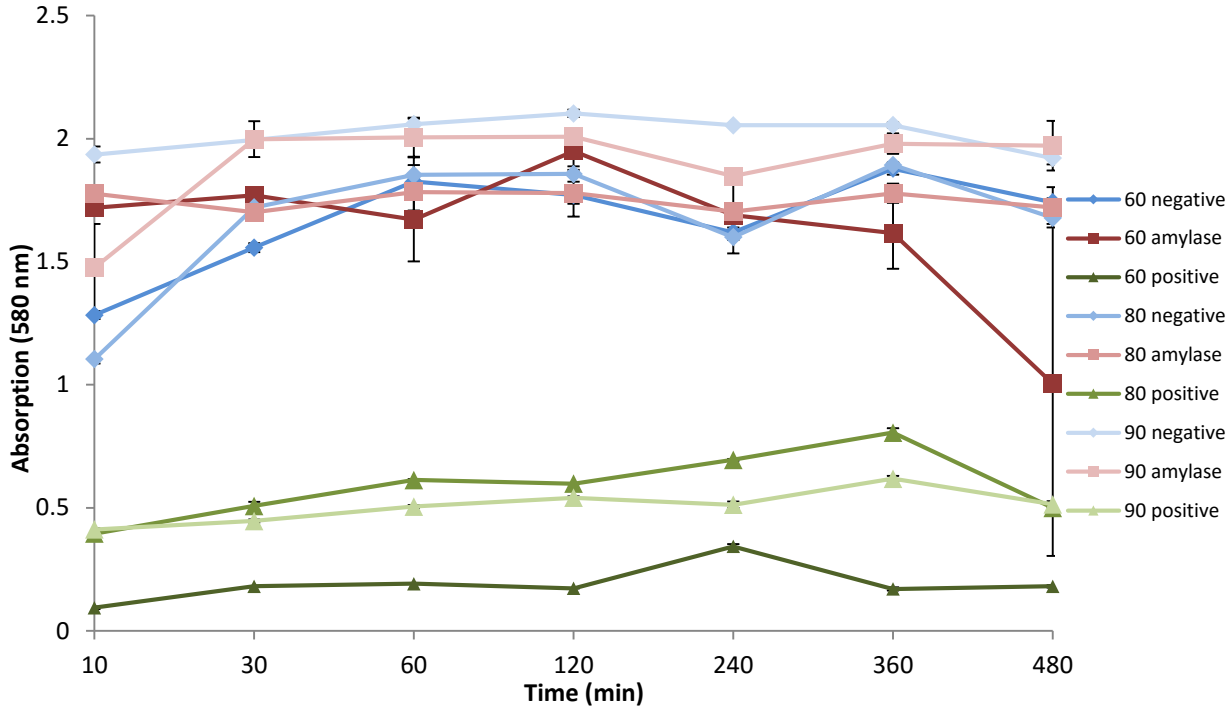


Figure 4.10 Temperature range of alpha-amylase (60 to 90 °C).

Alpha-amylase isolated from *B. coagulans* and expressed using *E. coli* BL21 Star and pET30a expression system had been Ni-NTA purified, then desalted so that the enzyme could be characterised and the temperature optimum could be found. Activity was assessed via optical density at 580 nm after 7 hours incubation at 60 °C, 80 °C and 90 °C in pH 8.78 buffer (containing 1% starch). Error bars show the standard deviation, n=4 for alpha-amylase and n=2 for negative control and positive control. The lower the absorbance the more activity was present as the iodine could only bind to remaining long chain sugars. *Negative control (♦); B. coagulans alpha-amylase (■); positive control (1 mg/mL Fisher bacterial alpha-amylase)(▲).*

4.2.7.5 Thermostability

B. coagulans alpha-amylase was assessed for thermostablity and by means of measuring activity after incubation at 50-100 °C for 30 min, then assayed at optimal conditions for 2 hours (section 2.2.6.14). Figure 4.11 shows that the alpha-amylase was not thermostable. However it could be seen that the positive control was stable up to 80 °C. Starch was broken down by the positive control meaning that the iodine was unable to bind to long chain sugars thus a lower optical density reading was obtained.

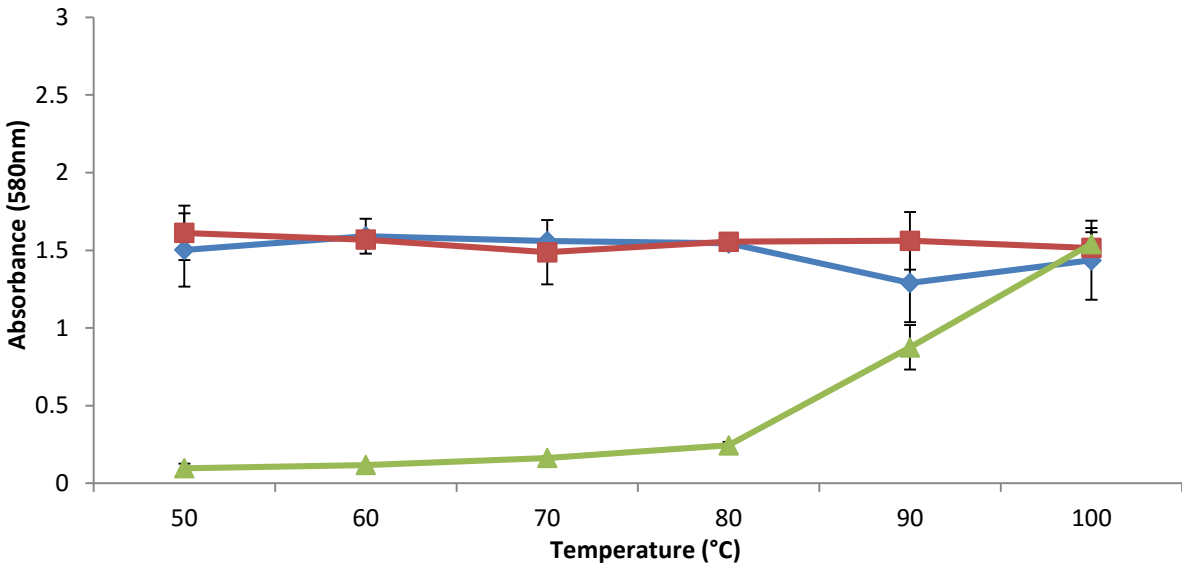


Figure 4.11 Thermostability of alph-amylase.

Alpha-amylase isolated from *B. coagulans* and expressed using *E. coli* BL21 Star and pET30a expression system had been Ni-NTA purified, then desalted so that the enzyme could be characterised and the thermostability could be shown. Activity was assessed via optical density at 580 nm after 2 hours incubation at 50 °C in pH 8.78 buffer (containing 1% starch). Error bars show the standard deviation, n=4. The lower the absorbance the more activity was present as the iodine could only bind to remaining long chain sugars. *Negative control (♦); B. coagulans alpha-amylase (■); positive control (1 mg/mL Fisher bacterial alpha-amylase)(▲).*

4.2.7.6 Cofactor Concentration

Cofactors have been shown to increase or inhibit the activity of many enzymes, in the case of alpha-amylase the cofactor was calcium (Feller *et al.*, 1992; Ramesh & Lonsane, 1989). Calcium chloride was used in a range of 0 mM-8 mM (section 2.2.6.15), based on findings in the literature, across a temperature range of 40-60 °C (Figure 4.12). The temperature range was used to establish if the addition of the cofactor enhanced the enzymes activity above and below the recognised optimum. The lower the absorbance the greater enzymatic activity was present due to fewer long chain sugars being available to bind to the iodine. Calcium concentration did not appear to have a major effect on the enzymatic activity. Activity also fluctuated at different temperatures. As Figure 4.12 shows, there was little change in activity over the calcium concentration at

40 °C, there was a very small amount of activity, yet this is unlikely to be as a result of the cofactor. With the increase of temperature to 50 °C there was also an increase in activity of the *B. coagulans* alpha-amylase, which rose with the increase of calcium until 3 mM whereby it plateaued. At 60 °C activity increased further with the calcium concentration becoming more important for enzyme activity, though 4 mM $CaCl^{2+}$ appeared to be the optimum.

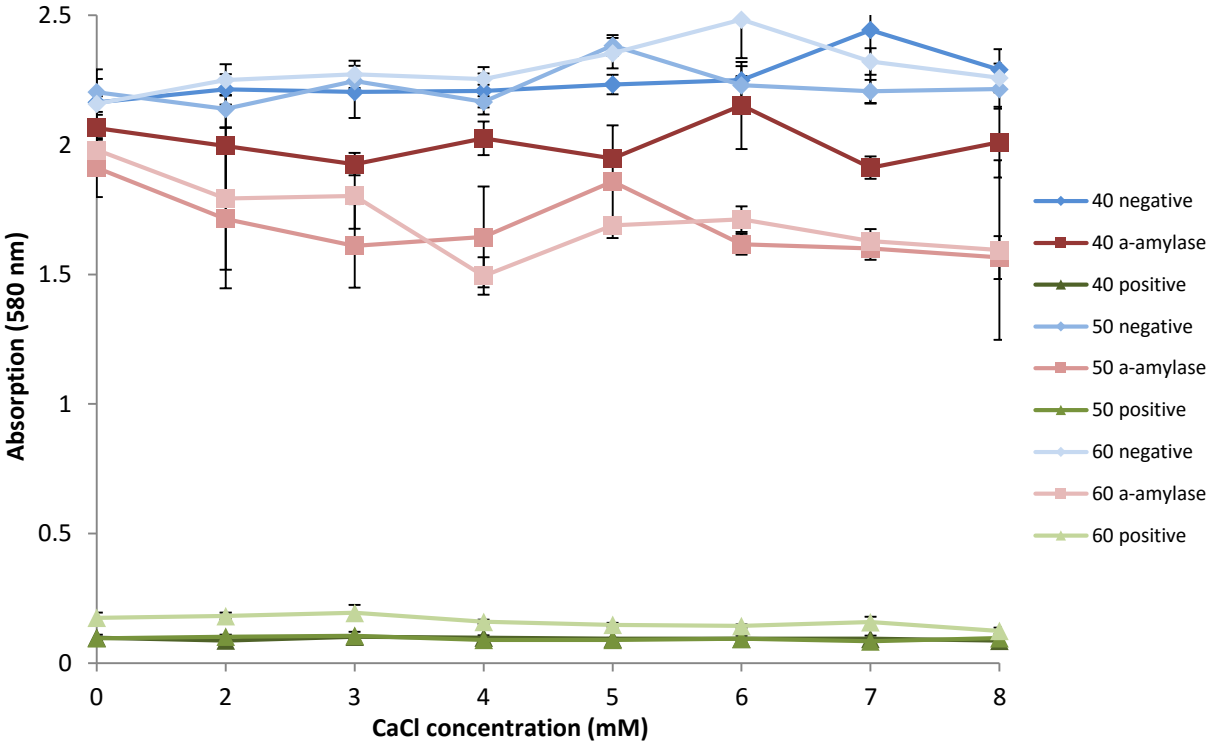


Figure 4.12 Cofactor concentration affect on alpha-amylase.

Alpha-amylase isolated from *B. coagulans* and expressed using *E. coli* BL21 Star and pET30a expression system had been Ni-NTA purified, then desalted so that the enzyme could be characterised and the effect of CaCl as a cofactor could be found. Activity was assessed via optical density at 580 nm after 4 hours incubation at 40 °C, 50 °C and 60 °C in pH 8.78 buffer (containing 1% starch). Error bars show the standard deviation, n=4. The lower the absorbance the more activity was present as the iodine could only bind to remaining long chain sugars. *Negative control (♦); B. coagulans alpha-amylase (■); positive control (1 mg/mL Fisher bacterial alpha-amylase)(▲).*

4.2.8 Endoglucanase discovery

A number of cellulose degrading enzymes which were secreted into the supernatant during isolation cultivation in 1% CMC broth were investigated. Proteins from the supernatant of 10 isolates that had activity at high pH and temperature (section 3.2.5.2) are presented in Figure 4.13. Some isolates were shown to have an additional band at 42.4 kDa (Figure 4.13), which were taken forward for zymography. However, only two samples, *B. subtilis* AHU and *B. licheniformis* AIB, had activity halos (Figure 4.16). Figure 4.14 shows the activity of ammonium sulphate precipitated proteins on simple plate assays which contained 1% CMC. *B. licheniformis* AJK had the most activity, presenting the largest halo, which peaked at pH 8. Only *B. licheniformis* AJG and *B. licheniformis* AJT had activity peaks at pH 7, the rest of the samples all peaked at pH 8. The positive control (Novozymes cellulase, NS22086) had a much lower pH optimum with most of its activity occurring at pH 5 although *B. licheniformis* AJK was the only sample to have more activity than the positive control at pH 8.

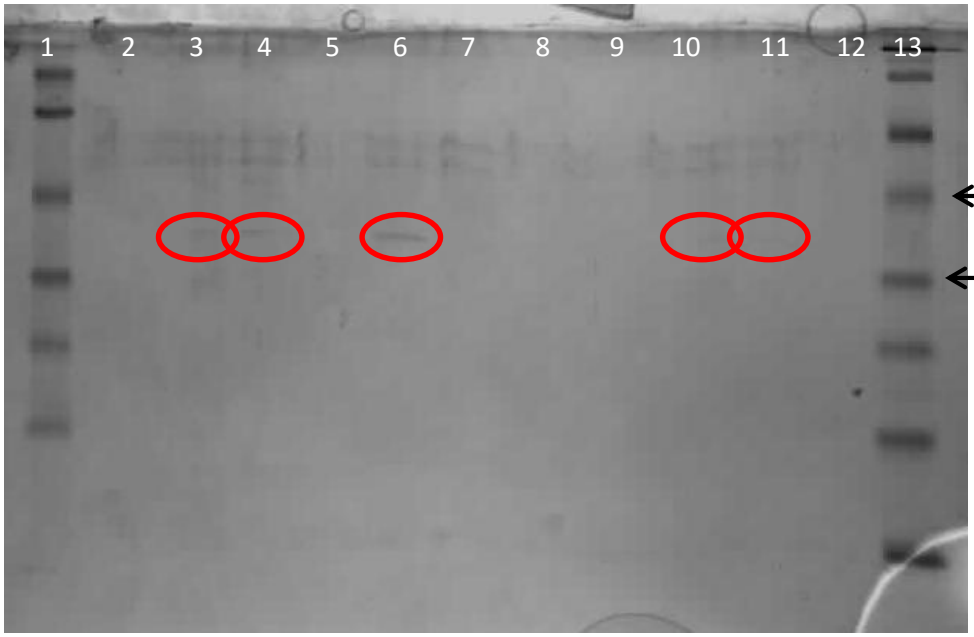


Figure 4.13 Coomassie stained 12% SDS-PAGE showing 80% ammonium sulphate precipitated crude protein.

From 1% CMC broth culture supernatant. Red ellipses show a band of 42.4 kDa in five samples. *1 – low range molecular marker (BioRad); 2 – boiled 1 µL Novozymes cellulase (NS22086) in 5 µL phosphate buffer (pH 7.5); 3 - B. subtilis AHU supernatant; 4 - B. licheniformis AIB supernatant; 5 - B. subtilis AIL supernatant; 6 - B. subtilis AJF2 supernatant; 7 - B. licheniformis AJG supernatant; 8 - B. subtilis AJH1 supernatant; 9 - B. licheniformis AJK supernatant; 10 - B. subtilis AJN1a supernatant; 11 - B. subtilis AJN1b supernatant; 12 - B. licheniformis AJT supernatant; 13 – broad range molecular marker (BioRad).* Photograph courtesy of Romain Jugal MSc, unpublished work.

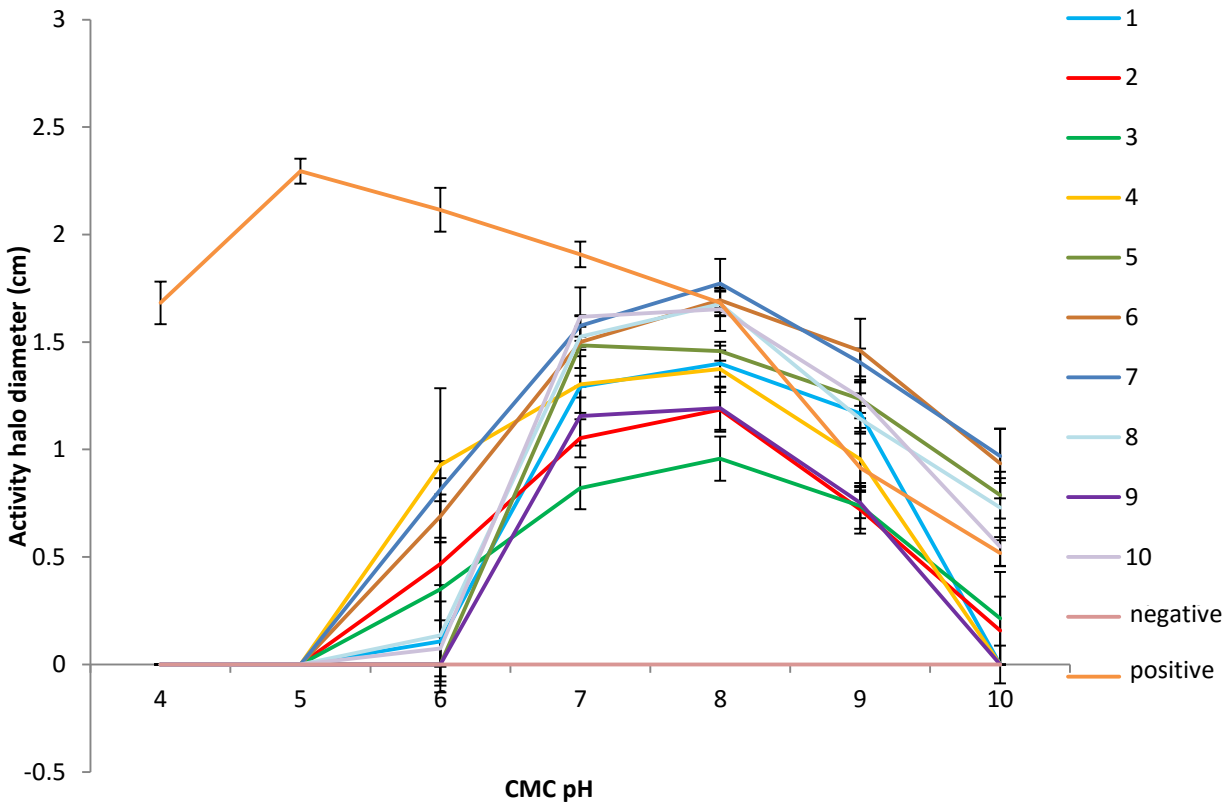


Figure 4.14 Activity of crude enzymes isolated from supernatant of cultures growing in 1% CMC broth on simple 1% CMC agar plate assay.

Proteins had been precipitated from the supernatant using 80% ammonium sulphate and were prepared fresh before the assay so that little ammonium sulphate would remain. Error bars show standard error, n=3. *1 – Bacillus subtilis AHU; 2 – Bacillus licheniformis AIB; 3 – B. subtilis AIL; 4 – B. subtilis AJF2; 5 – B. licheniformis AJG; 6 – B. subtilis AJH1; 7 – B. licheniformis AJK; 8 – B. subtilis AJN1a; 9 – B. licheniformis AJN1b; 10 – B. licheniformis AJT; négatif – negative control boiled 1 µL Novozymes cellulase (NS22086) in 5 µL phosphate buffer (pH 7.5); positif – positive control 1 µL (1000 BHU (2)/g) Novozymes cellulase (NS22086) in 5 µL phosphate buffer (pH 7.5).* Courtesy of Romain Jugal MSc, unpublished work.

Figure 4.16 shows preliminary tests with zymography using Native-PAGE containing 1% CMC, and though not very clear samples did have a small amount of activity near the top of the gel (red ellipse Figure 4.16 B)). The gel was split down the middle (through the negative control so that it was present on both gels) before coomassie staining for A) and further treatment for B). Modified native-PAGE gel A) (section 2.2.6.4) was coomassie stained to display the migration of proteins. Though visible proteins did not migrate very far down the gel they were calculated to be 42.4 kDa in size, based on

protein migration standard curve (Figure 4.15). The halos in Figure 4.16 B) were thought to be due to the 42.4 kDa proteins.

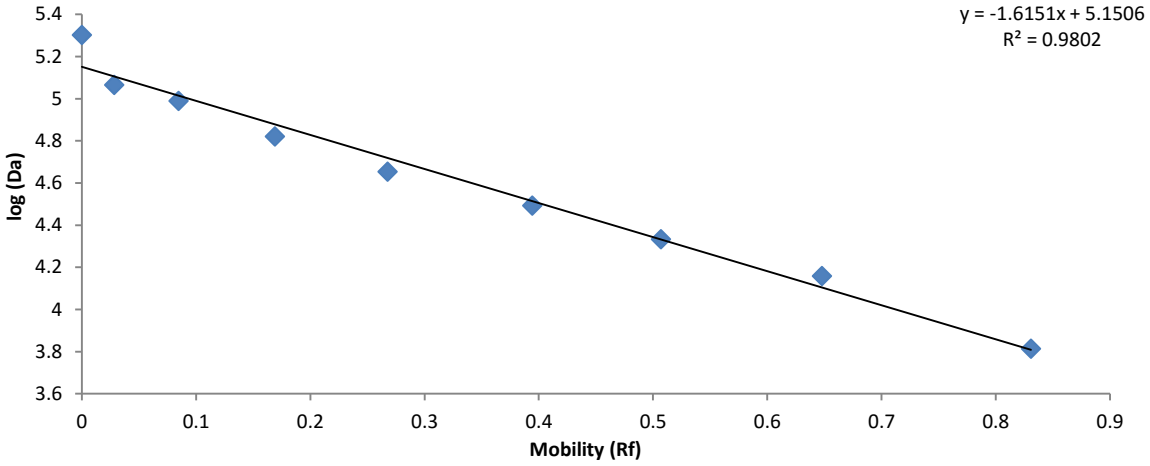


Figure 4.15 Standard curve of protein migration on 12% SDS-PAGE, Figure 4.13.

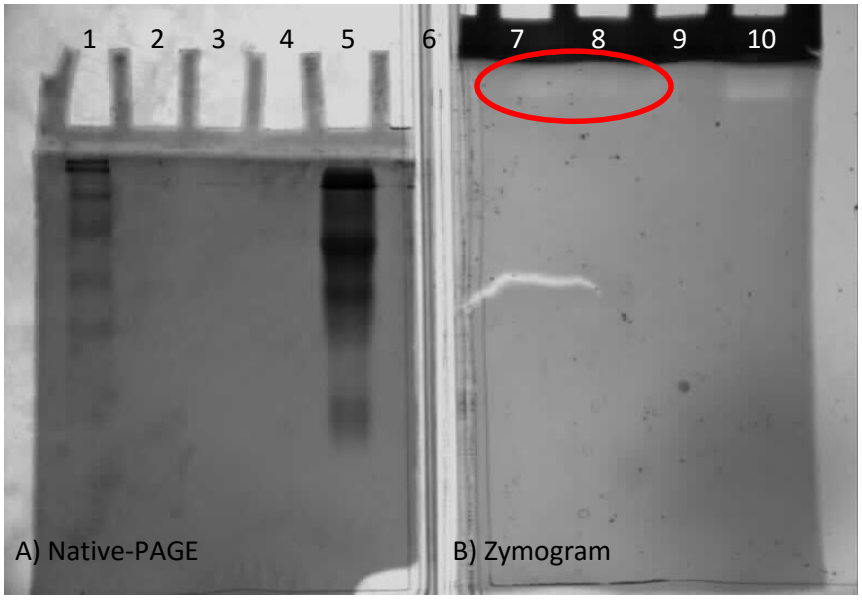


Figure 4.16 Zymography gel (1% CMC) of crude protein obtained via ammonium sulphate precipitation of culture supernatant showing cellulose degrading activity.

Red ellipse shows slight halos for samples *Bacillus subtilis* AHU and *Bacillus licheniformis* AIB on the zymography gel (B). *1 – low molecular marker (BioRad); 2 – 80% ammonium sulphate precipitated crude protein from B. subtilis AHU culture in 1% CMC broth; 3 - 80% ammonium sulphate precipitated crude protein from B. licheniformis AIB culture in 1% CMC broth; 4 - 80% ammonium sulphate precipitated crude protein from B. subtilis AHU culture in 1% xylan broth; 5 – positive control of 1 µL (1000 BHU (2)/g) Novozymes cellulase (NS22086) in 5 µL phosphate buffer (pH 7.5); 6 – negative control of boiled 1 µL Novozymes cellulase (NS22086) in 5 µL phosphate buffer (pH 7.5).* Photograph courtesy of Romain Jugal MSc, unpublished work.

4.3 Discussion

The aim was to express a gene encoding a useful enzyme which could potentially be used to depolymerise plant material in a biorefining context. Therefore, an alpha-amylase recombinant fusion protein was produced and characterised. *B. coagulans* ABQ was used due to its isolation conditions, as it withstood pasteurisation, and cultivation at pH 5.5 and 40 °C. Variants of the pET30a expression system detailed in this chapter have been used extensively in the literature (Ballschmiter, Armbrecht, Ivanova, Antranikian, & Liebl, 2005; Gírio *et al.*, 2010; Invitrogen, 2010; Jeng *et al.*, 2011), including the utilisation of the pGEM-T easy vector for the first stage of cloning (D. Y. Kim *et al.*, 2011). The cloning and expression process of obtaining protein was a successful and efficient method to obtain large amounts of near purified enzyme for further characterisation. However, when attempting to obtain xylan deacetylase (also from *B. coagulans* ABQ) for ligation and transformation, no positive results were achieved. Xylan deacetylase has not been shown to be toxic for cloning so it is unknown why the gene would not transform correctly into the first vector. However, a cloning system has been detailed whereby pETBlue2 could be used with *E. coli* BL21 as both cloning and expression system, which allows for the cloning of toxic or difficult genes (L. E. Taylor, Weiner, Hutcheson, Ekborg, & Howard, 2014). Such a system could be used in the future.

Production of the cloned recombinant fusion alpha-amylase protein appeared to be “leaky” as the enzyme may have been present in inclusion bodies (Henco, 2003; Stampolidis, Kaderbhai, & Kaderbhai, 2009). The alpha-amylase was present in the pellet and supernatant after freeze/thaw, lysozyme and sonication treatment thus the

protein was both soluble and insoluble (Stampolidis *et al.*, 2009). However, the protein yield was high enough for use in the characterisation assays. Aggregation of the proteins could have produced the two expressed bands on the gels (Bondos & Bicknell, 2003), however this is a common problem and can be avoided by using DTT instead of β-mercaptoethanol in the sample buffer (F. Sun, 2015). Figure 4.8 shows two pH optimas of the alpha-amylase isolated from *B. coagulans* ABQ. It has been documented in the literature that enzymes are able to have two peaks of activity (Ferkovich & Mayer, 1975; Holzman, Fesik, Park, & Kofron, 1991; Horikoshi, 1996; Merritt & Karn, 1977; Sticher & Jones, 1992; Thompson, 1993) so it was not unusual for the *B. coagulans* alpha-amylase to exhibit this.

Figure 4.5 shows that the simple starch assay was a very visual and effective screening assay which supported the zymogram shown in Figure 4.4 but was able to show the intensity of each sample. Zymography was useful to demonstrate that both expressed bands (Figure 4.3) were active in degrading starch and thus most likely to be the alpha-amylase protein. The smaller band (56 kDa) was expected as calculated from the SDS-PAGE gel (4.2.3) and amino acid sequence using the Protein Calculator (section 2.2.7.9). Western Blot transfer of proteins using the His-tag showed that both bands were transferred and detected on the Western blot. If the enzyme had renatured when using SDS-PAGE zymography, the size could have been determined for both bands using the protein marker. Instead, the molecular weight for the other alpha amylase (section 5.3.5.3) was calculated (via amino acid sequence) using the Protein Calculator. This alpha amylase was calculated to be 99.78 kDA, which could be the correct size of the band on the gel (Figure 4.3), though the amino acid sequence had

very little homology to the expressed amylase (Appendix 3: 8.3.14). After much troubleshooting, modified Native-PAGE was used, unlike conventional Native-PAGE (Reisinger & Eichacker, 2007; Upadhyay *et al.*, 2005; Wittig, Braun, & Schagger, 2006; Wittig & Schägger, 2005). Therefore, the protein marker and samples did not separate into linear bands due to denaturation as in SDS-PAGE. However, it was unknown why the larger band was present.

The most likely explanation for two bands appearing on the gel after expression and purification (Figure 4.3) is protein dimerization (Ballschmiter *et al.*, 2005). Maltogenic-amylases have been shown to dimerise (Cho *et al.*, 2000; Mehta & Satyanarayana, 2013), and such dimers were proven to be more thermostable (Mehta & Satyanarayana, 2013). Those that had a monomer-dimer equilibrium were stabilised as a dimer, in the presence of zinc (Tellam, Winzor, & Nichol, 1978). Alternatively, the larger band could have been the other alpha-amylase, previously mentioned. However, this should have been seen on the gel electrophoresis as three bands. Also, the primers were designed for the start and end of the alpha-amylase gene, which was dissimilar to the alpha-amylase targeted (alpha amylase 1 section 5.3.5.3). It has been shown that alpha-amylases can be active in both monomeric and dimeric forms (Dong, Vieille, Savchenko, & Zeikus, 1997), thus this is the most likely explanation. Alternatively, the larger band could have been the alpha-amylase protein aggregated (Bondos & Bicknell, 2003; Stampolidis *et al.*, 2009).

The alpha-amylase enzyme obtained via cloning and expression would need to be purified further so that only the alpha-amylase was present in each sample. In an attempt to gain pure protein, a cell free approach was taken using Bioneer ExiProgen

EC1 protein synthesis kit (Bioneer Corp., 2014) using the recombinant fusion alpha-amylase. However, the procedure was unsuccessful. After a series of vector concentrations were assessed the quantity of vector (with alpha-amylase gene insert) required to gain a protein product was calculated to be 6 μg or higher. Another problem encountered was a preliminary HPLC run looking for interesting products from cellulose and xylose metabolism from a selection of isolates (Figure 4.14). A large glucose peak was obtained in all results (including those grown on xylan) with few other peaks, making the readings unreliable. The column was changed however there was not enough time to run the samples again. Also the Phadebas assay (Figure 4.6) calculations may not be accurate according to the standard curve (Appendix 8.2.1) as the assay was completed at a different temperature.

Proteins from culture supernatant were precipitating using 80% ammonium sulphate (section 2.2.6.8) which did not denature enzymes unlike TCA (section 2.2.6.7). Proteins were assessed using 12% SDS-PAGE to view any secreted enzymes (Figure 4.13). Proteins were then analysed using Native-PAGE gels containing 1% CMC for zymography (Figure 4.16) to ascertain whether the bands shown in Figure 4.13 were endoglucanase enzymes. Halos of activity on the zymogram and simple plate assays (Figure 4.14) demonstrated that cellulases were most likely secreted into the media by isolates, to degrade the long chain sugars, before transporting monomers or polymers into the cell for metabolism (Simonen & Palva, 1993). Unfortunately, the xylanolytic zymography was unsuccessful; though proteins were secreted the zymograms and simple plate assays did not show consistent results.

4.3.1 Conclusions

Though the methods used for the experiments in this chapter were efficient and had a good success rate for alpha-amylase production, there were problems when trying to express other genes, such as xylan deactylase. Therefore more troubleshooting and optimisation would have to be utilised for overall effectiveness. Also these techniques only utilised known or partially known genes, whereas SDS-PAGE and zymography screens for secreted enzymes could identify completely novel genes, which could be sequenced using mass-spectrometry. Metabolic products could also be identified using HPLC or mass-spectrometry. In future a variety of enzymes could be isolated and characterised using both techniques to isolate novel genes. With such a plethora of literature available on this bacterium (*B. coagulans*) it was interesting to see if we could find anything different with protein isolation and genome sequencing (sections 5.3.2 and 5.3.5.3).

Chapter 5 Next Generation Sequencing of *Miscanthus* Associated Bacteria and Protein Modelling of Sugar Releasing Enzymes

5.1 Introduction

Genome sequencing followed by enzyme mining has been widely discussed and is well established for the discovery of novel enzymes (Kennedy, Margassery, Morrissey, O'Gara, & Dobson, 2013; Mutondo, Huddy, Bauer, Tuffin, & Cowan, 2012). When enzyme mining through the metagenome, little is known about the natural host (Berlemont & Galleni, 2012). Whole genome sequencing from cultured isolates eradicated this problem.

Xylan and cellulose have been reported as abundant renewable fermentable feedstocks (Harish, Janaki Ramaiah, & Babu Uppuluri, 2015; Wooley *et al.*, 1999). Starch hydrolysis has also been shown to be important for food, pharmaceutical and chemical industries (Berlemont & Galleni, 2012). Enzymes able to degrade cellulose, xylan and starch generally require resistance to high temperatures for industrial application due to substrate insolubility (van den Burg, 2003). Thus these polymers, and some elevated temperatures, were the focus of study for carbon utilisation assays identified in Chapter 3 and enzyme identification and modelling described in this chapter. Therefore, it was hypothesised that bacteria exhibiting extremophile tendencies would have enzymes of use in biorefining. Also, there would be potential to find novel or adapted enzymes.

Nextera XT library preparation included 23 isolates that were chosen for next generation sequencing (section 2.2.7). The isolates were chosen based on their ability to grow and had activity at a range of pHs and temperatures shown in Chapter 3; specifically due to their ability to depolymerise xylan at extremes of pH. The 23

genomes encompassed eight different bacterial species based on 16S rRNA amplicon sequencing (section 3.2.1). There were nine *Bacillus licheniformis*; four *Bacillus subtilis*; two *Bacillus pumilus*, one *Bacillus coagulans*; one *Brevibacillus brevis*; one *Burkholderia fungorum*; one *Burkholderia sp*., and four *Unknown sp*..

Following sequencing a range of bioinformatics techniques were used to analyse the sequences. Sequence files were first imported into CLC Genomics for contig assembly and preliminary annotation (section 2.2.7.4). Secondly, contigs were uploaded to RAST for automatic annotation. Subsystems were analysed using RAST Seed Viewer (section 2.2.7.5) and enzymes identified in active pathways for protein modelling. Modelling took place using Phyre2 (section 2.2.7.6) and 3DLigandSite (section 2.2.7.7).

5.2 Results: Genomic libraries

Bacterial isolates were selected based on their ability to degrade sugars at extremes of pH and temperature shown in Chapter 3 via carbon utilisation assays. Genomic DNA of selected isolates was used to construct barcoded Nextera XT libraries (Materials and Methods 2.2.7). Libraries were sequenced on the Illumina MiSeq platform using paired end and sequencing, with the view to isolate genes and pathways related to sugar degradation. Quantification of gDNA from 1 mL cultures occurred via Qubit as this was more accurate and sensitive than the Epoch NanoDrop (BioTek Instruments, Inc. Winooski) method. Genomic DNA from each of the selected isolates was diluted to 1 ng/μL in a 96 well plate format, whereby it could be diluted to 0.2 ng/μL as required by the Nextera XT kit. Dilutions took place based on initial Quibit readings. Libraries were then quantified again using Qubit after PCR amplification (Materials and Methods 2.2.7.3.3). Samples were checked for quality prior to Illumina sequencing.

5.3 Results: Genomic analysis

Bacterial genomes of 23 isolates were paired end sequenced via Illumina MiSeq (section 5.2). Reads were imported into CLC Genomics Workbench, trimmed, merged and contigs constructed (section 2.2.7.4). Contigs were uploaded to RAST and automatically annotated (section 2.2.7.5). Genome sizes ranged from 3.1 Mbp to over 13 Mbp, and GC content ranged from 41.2-65.2% (Table 5.1). Coverage of actual genome size ranged from 9.5-559x.

16S rRNA genes were found within the genome and aligned using NCBI BLAST to assess original identification. Sequence based comparisons to closest neighbours were also used to validate identifications (section 5.3.1). All isolates had a single copy of the 16S rRNA gene bar *Burkholderia sp.* AEW2a which had two copies of the 16S rRNA gene, located on contigs 291 and 203. Two versions of this gene could show that there were two species of bacteria/*Burkholderia* present. As the total genome size was over 13 Mbp (Table 5.1) it was most likely that two or more genomes were present. Previously thought mixed isolates AFDa (closest neighbours *Dyella japonica* and *Rhodanobacter spathiphylli*) and AEOb (closest neighbour *Rhodanobacter spathiphylli* B39) also had one copy of the 16S rRNA gene. A single copy of the 16S rRNA gene could indicate that these genomes are singular rather than mixed. However, sections of the genome may not have been amplified and therefore missed in sequencing. Alternatively, the 16S gene may have been too variable to align to the closest neighbour gene and could therefore be present but overlooked.

Table 5.1 Illumina MiSeq read and contig information for each bacterial genome sequenced.

From both CLC Genomics Workbench 7.03 and SEED RAST. *Burkholderia sp.* AEW2a had a total genome size of 13 982 873 bp and was therefore most likely to be two or more genomes combined. ** - isolates believed to be mixed colonies; ^ - previously "Unknown sp." (could not be identified by 16S rRNA ribotypying).*

Name	No. reads	Average length of reads (bp)	Total bp	Coverage of est. genome size	GC content (%)	No. contigs in draft genome	Average contig size (bp)	Total size of contigs combined (bp)	Coding sequences	Possibly missing genes	No. subsytems
B. coagulans ABQ	13 729 916	143.5	1 932 716 813	559x	46.3	151	22 887	3 455 951	3 868	12	434
B. licheniformis AHY1	333 155	204.28	68 056 281	15.3x	45.7	270	16 573	4 441 645	4 799	35	489
B. licheniformis AIB	700 750	230.16	161 285 259	39.6x	43.7	125	33 703	4 077 933	4 182	14	476
B. licheniformis AJF1	1 177 808	207.7	244 631 231	78.9x	46.2	124	34 531	3 101 910	3 332	19	383
B. licheniformis AJG	1 552 391	202.61	314 527 087	72.4x	45.9	150	29 167	4 345 800	4 533	13	486
B. licheniformis AJH1	1 556 359	210.04	326 892 328	74.3x	45.8	163	28 038	4 401 586	4 769	26	490
B. licheniformis AJN1b	659 971	187.47	123 723 761	29.4x	46.2	130	32 841	4 203 665	4 524	27	485
B. licheniformis AJT	751 763	182.39	137 114 136	32x	45.8	477	9 091	4 291 006	4 568	27	485
B. licheniformis AJVa	2 012 651	212.99	428 861 476	98.1x	45.8	197	22 764	4 370 194	4 675	36	489
B. pumilus AIF	625 542	202.68	126 787 136	33.1x	41.2	182	22 635	3 824 707	4 017	16	471
B. pumilus AIK	532 021	109.92	101 572 511	26.8x	41.1	158	25 810	3 793 583	3 945	2	473
B. subtilis AHU	255 364	187.22	47 808 388	10.8x	43.3	794	5 558	4 408 074	4 550	109	463
B. subtilis AJF2	282 932	180.92	51 187 018	12.4x	43.8	457	8 997	4 112 085	4 283	47	470
B. subtilis AJK	1 304 990	215.23	282 689 360	65.4x	43.3	209	21 199	4 324 397	4 537	11	476
B. subtilis AJN1a	371 803	176.22	65 518 777	15x	43.7	602	7 278	4 366 399	4 507	47	468
Brevibacillus brevis AHW	569 943	199.78	113 866 043	20.2x	54.1	401	14 034	5 627 994	5 561	71	462
Burkholderia fungorum AELa	796 119	203.99	162 400 234	18x	61.8	529	18 115	9 038 305	8 483	101	527
Burkholderia sp. AEW2a*	1 332 238	200.32	266 868 778	19.1x	62.8	1 710	8 364	13 982 873	12 867	220	579
Shingomonas	3 424 239	211.65	724 742 139	138.6x	65.2	149	10 486	5 230 456	4 804	62	439

asaccharolytica AEH1											
Luteibacter rhizovicinus AEOb*^	302 036	151.07	45 627 791	9.5x	64.6	1 068	4 479	4 784 038	4 204	119	428
L. rhizovicinus AEZa^	433 081	201.61	87 314 491	18x	64.6	376	14 372	4 857 043	4 301	79	443
L. rhizovicinus AFDa*^	358 816	195.33	70 088 986	14.4x	64.5	455	11 118	4 858 622	4 292	99	434
B. licheniformis AIE1^	429 992	202.21	86 948 782	19.2x	45.4	307	14 878	4 537 679	4 883	28	489

5.3.1 Whole genome comparisons using Seed-RAST

Sequence based comparisons of genomes with "closest strains" (Materials and Methods 2.2.7.5) within RAST database was performed for each isolate. *Burkholderia fungorum* AELa had sections of low homology of 10% with *Burkholderia xenovorans* LB400 (Figure 5.1a) but high homology to *Burkholderia fungorum* NBRC 102489 (99.5-100%) (Figure 5.1b). Differences in *Burkholderia fungorum* AELa genome to reference were mainly located on contig 10 of the reference. However, these genes had been lost from the AELa genome or added into the reference genome.

Bacillus licheniformis AJVa had very high homology (99-100%) with *Bacillus licheniformis* ATCC 14580, with some regions of no or little homology. Up to four genomes could be compared at the same time. Four *B. licheniformis* strains were compared to *B. licheniformis* ATCC 14580 (Figure 5.2). Genomes of the same species were compared to each other using this technique to view differences more specifically between strains than to a published reference. When a query strain was used as the reference genome the pattern displayed in Figure 5.2 altered. The query genomes were assembled into contigs only and therefore could not be ordered into an appropriate assembly when compared to another query strain.

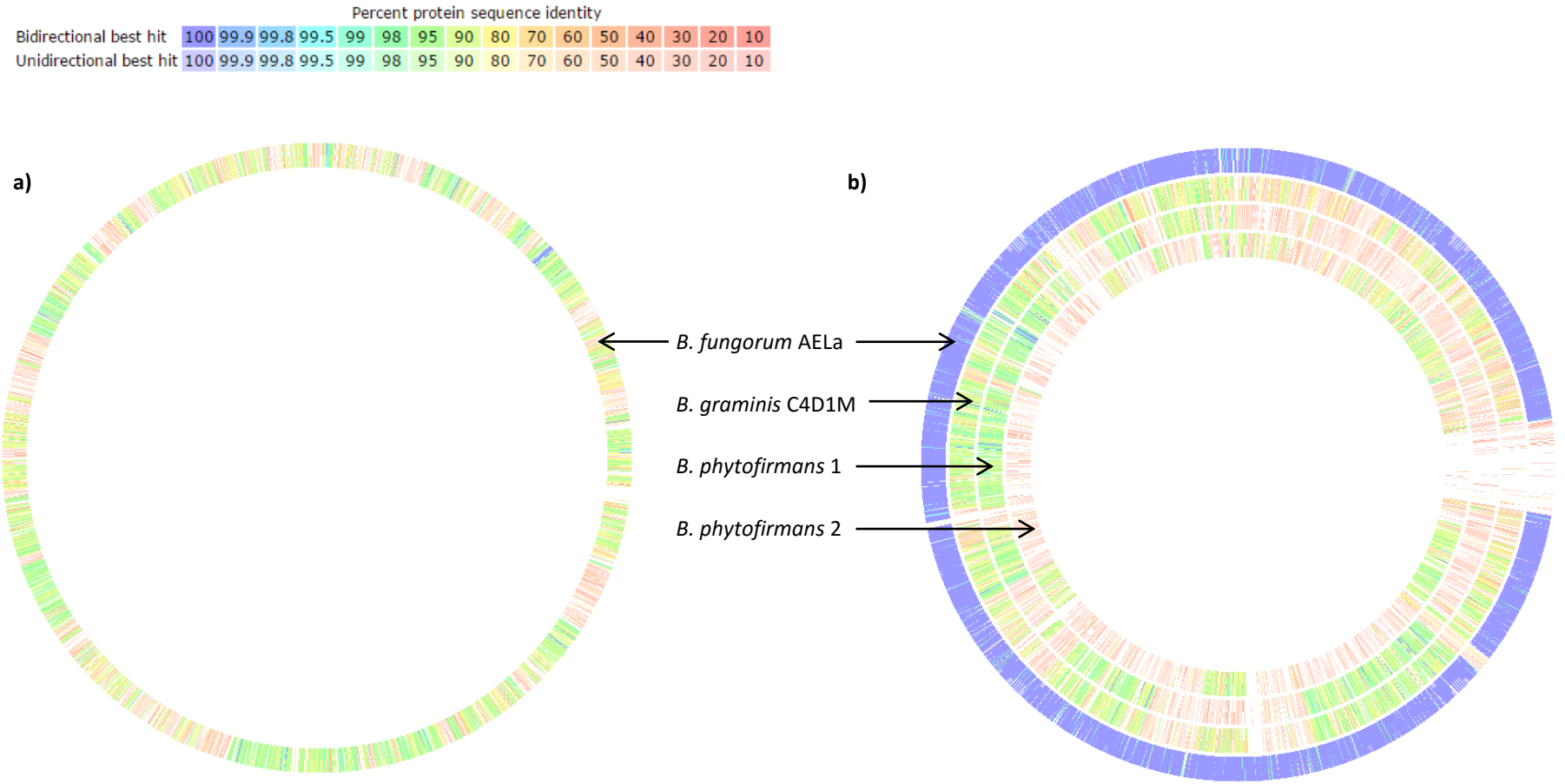


Figure 5.1 *Burkholderia fungorum* AELa compared

a) Reference strain *B. xenovorans*; b) reference strain *B. fungorum* NBRC 102489 and other *Burkholderia sp*. There was high homology between the two *B. fungorum* sequences, though a section of AELa genome was missing from sequencing.

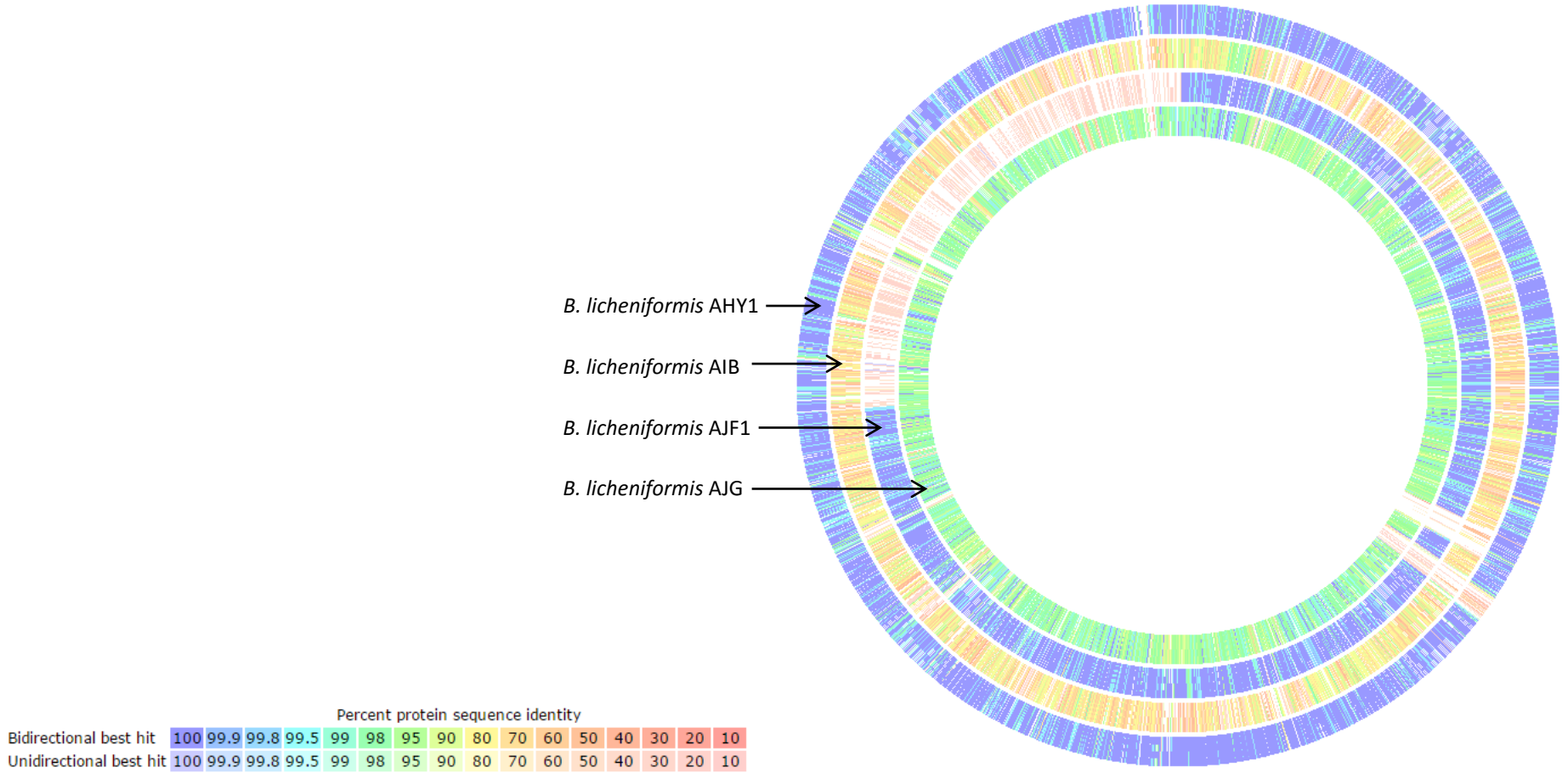


Figure 5.2 Comparison of four *B. licheniformis* (AHY1, AIB, AJF1 and AJG) strains against reference genome *B. licheniformis* ATCC 14580.

Strains were sequenced via Illumina MiSeq, contigs assembled in CLC Genomics and annotated using RAST Seed Viewer.

Bidirectional homology, whereby replication of genes can proceed in both directions and in differing frames (P. K. Gupta, 2014; Koyanagi, Hagiwara, Itoh, Gojobori, & Imanishi, 2005; Kunst *et al.*, 1997; van der Veen, Harris, O´Toole, & Claesson, 2014), was prevalent in whole genome comparisons. A range of homology was present between sequenced isolates (Figure 5.2). *B. licheniformis* AJF1 had high homology to the reference for most of the genome. However, a quarter of the genome had low homology (10-40%) which could be due to horizontal gene transfer. When compared to a different reference genome (*B. licheniformis* 9945A) homology was up to 98%, however the low homology gap remained. Though this section spanned a number of contigs, there were many gaps and any homology was unidirectional.

The closest neighbour function was utilised in an attempt to identify isolates classed as "*Unknown sp.*". However, the five genomes with this classification had little homology with species suggested (Figure 5.3a). AEOb and AEZa had had similar profiles across the carbon utilisation assays (Figure 3.10, Appendix Table 8.2). *Sphingopyxis alaskensis* RB2256 was recorded as closest neighbour to AEH1. *Dyella japonica* A8 was suggested to be closest neighbour to AFDa. However, AFDa had high homology to AEOb and AEZa, though AEH1 had very little homology (Figure 5.4). Alignment of the 16S rRNA genes present in each genome revealed that AEOb, AEZa and AFDa were all strains of *Luteibacter rhizovicinus*, whereas AEH1 was most likely *Sphingomonas asaccharolytica*. However, AIE1 had high homology to closest neighbour *B. licheniformis* ATCC 14580 and could therefore be classified as a *B. licheniformis* strain. However, there was a lower percentage of homology when compared with *B. licheniformis* 9945A (Figure 5.6).

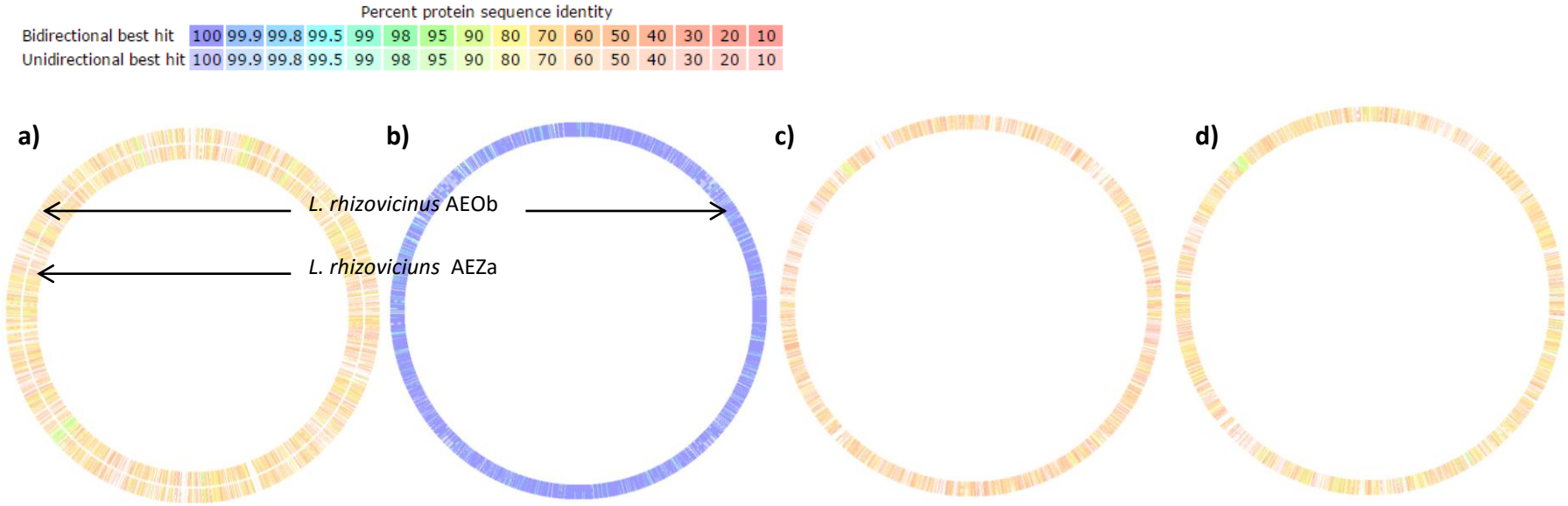


Figure 5.3 Genomes previously classed as *Unknown sp.* compared.

The *L. rhizovicinus* and *S. asaccharolytica* genomes were not available for comparison. a) AEOb and AEZa to closest neighbour *Rhodanobacter spathiphylli* B39. Homology was low (10-80%). b) *L. rhizoviciuns* AEOb (as the reference genome) and *L. rhizovicinus* AEZa. There was high homology (99.5-100%) between the two isolates. c) *S. asaccharolytica* AEH1, and *Sphingopyxis alaskensis* RB2256 as the reference genome. d) *L. rhizovicinus* AFDa, and *Dyella japonica* A8 as the reference genome. Low homology (10-80%) was observed for both *c* and *d* with their closest neighbours suggested by RAST.

Figure 5.4 Multi-genome comparison for previously "*Unknown sp.*" with *L. rhizovicinus* AEOb as the reference.

AEH1 had very little homology to AEOb, AEZa and AFDa, which all had very similar sequence based homology.

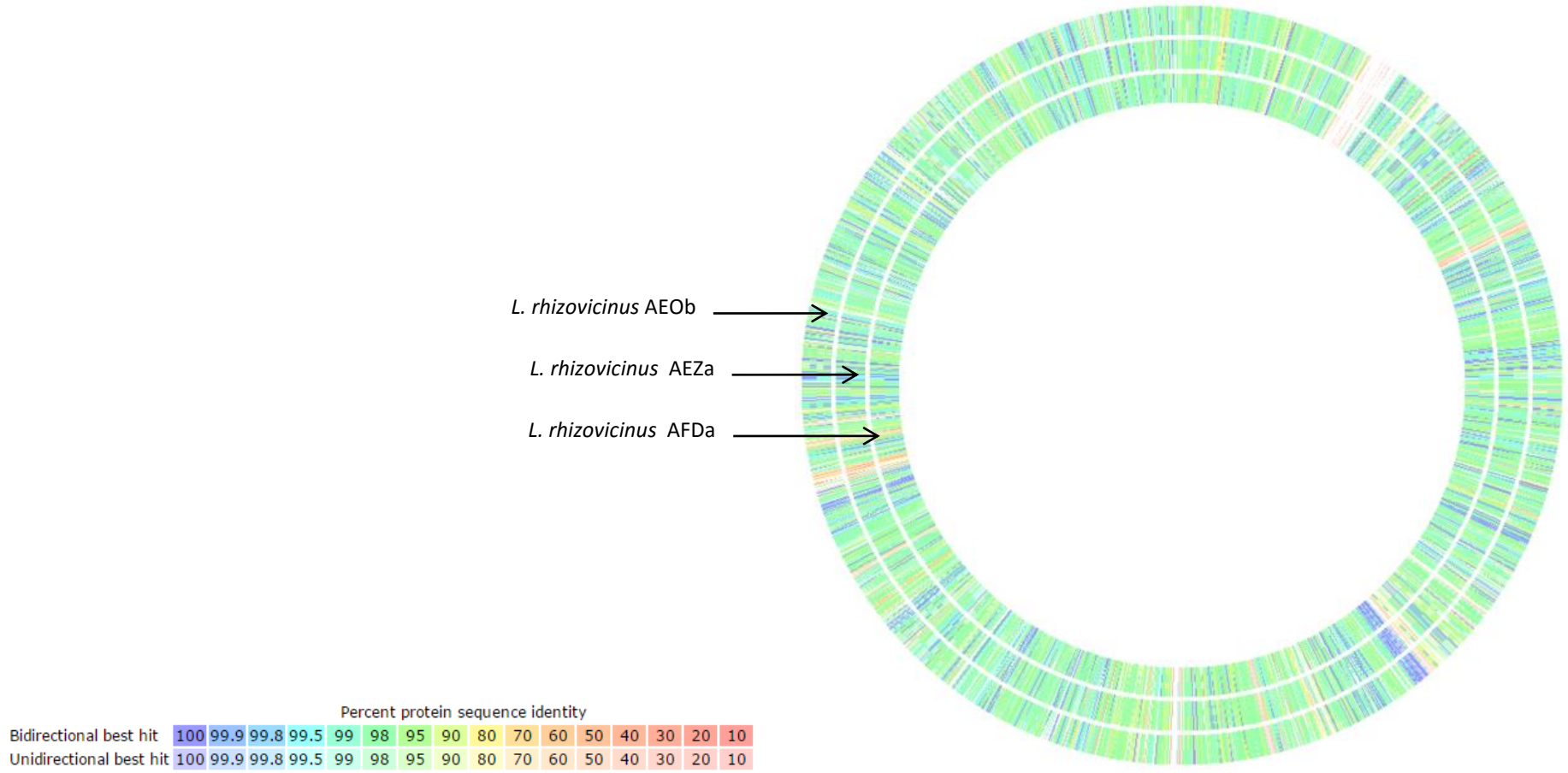


Figure 5.5 AEOb, AEZa and AFDa aligned with *L. rhizovicinus* DSM 16549.

Regions of high homology were present, though there was on average 95% homology across the genome. However, AEOb, AEZa and AFDa appeared to be almost exactly homologous.

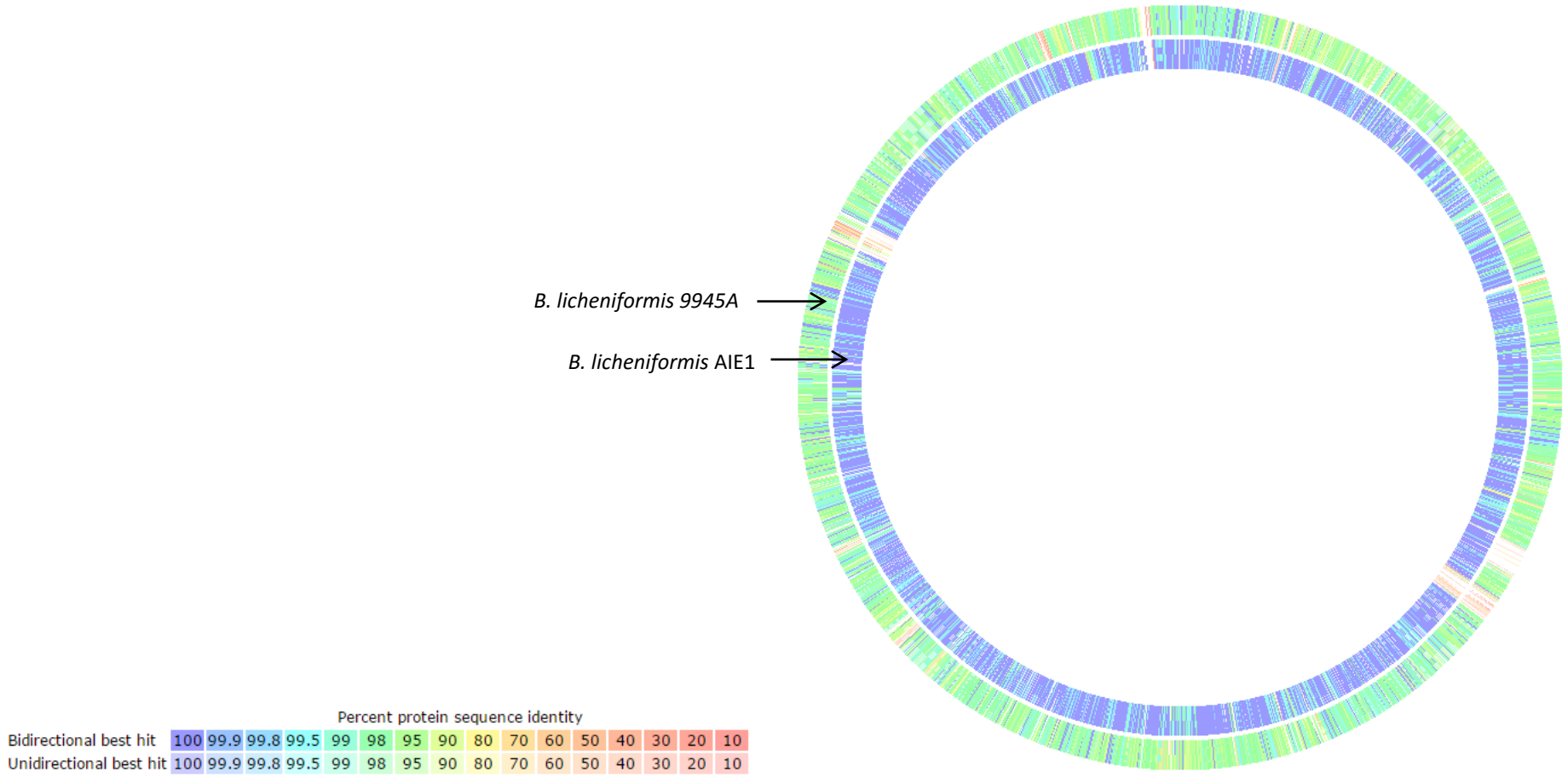


Figure 5.6 AIE1 compared with closest neighbour *B. licheniformis* ATCC 14580 and *B. licheniformis* 9945A.

AIE1 had high homology to ATCC 14580 (as the reference genome) which has low homology to 9945A.

Table 5.2 Subsystem coverage details for each query strain genome.

Genomes listed alphabetically. * - two genomes present in a single sample; ^ - less than 97% homology

16S rRNA identification	Genome	Subsystem			Non-Subsystem		
		Coverage (%)	Total genes	Hypothetical proteins	Coverage (%)	Total genes	Hypothetical proteins
B. coagulans	*B. coagulans* ABQ	45	1721	93	55	2147	1422
*B. licheniformis^*	*B. licheniformis* AHY1	45	2158	141	55	2641	1620
B. subtilis	*B. licheniformis* AIB	49	2037	128	51	2145	1171
Unknown sp.	*B. licheniformis* AIE1^	45	2164	148	55	2719	1707
B. licheniformis	*B. licheniformis* AJF1	42	1384	93	58	1948	1061
*B. licheniformis^*	*B. licheniformis* AJG	47	2120	136	53	2413	1401
B. subtilis	*B. licheniformis* AJH1	46	2155	145	54	2614	1602
B. licheniformis	*B. licheniformis* AJN1b	47	2096	139	53	2428	1439
*B. licheniformis^*	*B. licheniformis* AJT	47	2119	141	53	2449	1464
B. licheniformis	*B. licheniformis* AJVa	46	2137	138	54	2538	1546
*B. pumilus^*	*B. pumilus* AIF	47	1869	100	53	2148	1189
B. licheniformis	*B. pumilus* AIK	47	1852	102	53	2093	1160
B. subtilis	*B. subtilis* AHU	45	2038	127	55	2512	1450
B. subtilis	*B. subtilis* AJF2	48	2024	120	52	2259	1268
B. licheniformis	*B. subtilis* AJK	47	2088	129	53	2449	1419
B. subtilis	*B. subtilis* AJN1a	46	2048	125	54	2459	1415
B. agri	*B. brevis* AHW	42	2299	123	58	3262	1897
B. fungorum	*B. fungorum* AELa	44	3718	229	56	4765	2300
B. fungorum	*Burkholderia sp.* AEW2a*	45	5672	356	55	7195	3619
Unknown sp.	*S. assacharolytica* AEH1	43	2047	97	57	2757	1423
Unknown sp.	*L. rhizovicinus* AEOb	42	1761	102	58	2443	1233
Unknown sp.	*L. rhizovicinus* AEZa	44	1874	113	56	2427	1187
Unknown sp.	*L. rhizovicinus* AFDa	43	1831	108	57	2461	1196

5.3.2 Genomic analysis of *Bacillus coagulans* ABQ

A functional comparison of genomes was carried out using RAST. A subsystem was selected and presence/absence of genes compared to a reference genome (if available). This function was also used to find which genes were present in the query strain genome but not in the reference. Thus *Bacillus coagulans* ABQ (Chapter 4) was compared to the closest neighbour (*B. coagulans* 36D1) for genes present in ABQ but absent in 36D1 (Table 5.3). Functional comparison yielded 15 genes that were only present in *B. coagulans* ABQ. Alpha-glucosidase (EC 3.2.1.20) has been shown to degrade sucrose to β-D-fructose, and maltose to α-D-glucose (Figure 5.7). Interestingly the reference genome (36D1) did not have this enzyme present. Beta-glucosidase (EC 3.2.1.21) degrades cellobiose to β-D-glucose and glucoside to α-D-glucose (Figure 5.7). Again this enzyme was not present in the reference genome. Arginine decarboxylase/agmatinase/agmatine deiminase pathway was present in both *B. cogulans* ABQ and 36D1 genomes. However, *B. coagulans* ABQ also had three additional genes from the Arginine deiminase (ADI) pathway (Table 5.3), which is an anaerobic route (Zúñiga, Pérez, & González-Candelas, 2002). This is not unusual as *B. coagulans* are facultative anaerobes (De Vecchi & Drago, 2006).

Table 5.3 Functional comparison of *B. coagulans* ABQ with closest neighbour *B. coagulans* 36D1.

Genes selected for presence in ABQ and not in the comparison genome. Generated in RAST Seed Viewer (Overbeek *et al.*, 2014).

Category	Subcategory	Subsystem	Role	*Bacillus coagulans* ABQ Gene
Amino Acids and Derivatives	Arginine; urea cycle, polyamines	Arginine Deiminase Pathway	Arginine deiminase (EC 3.5.3.6)	fig|345219.18.peg.1936
Amino Acids and Derivatives	Arginine; urea cycle, polyamines	Arginine Deiminase Pathway	Carbamate kinase (EC 2.7.2.2)	fig|345219.18.peg.1939
Amino Acids and Derivatives	Arginine; urea cycle, polyamines	Arginine Deiminase Pathway	Transcriptional regulator ArcR essential for anaerobic expression of the ADI pathway, Crp/Fnr family	fig|345219.18.peg.1940
Carbohydrates	Di- and oligosaccharides	Beta-Glucoside Metabolism	Beta-glucosidase (EC 3.2.1.21)	fig|345219.18.peg.1177
Carbohydrates	Monosaccharides	D-Galacturonate and D-Glucuronate Utilization	Alpha-glucosidase (EC 3.2.1.20)	fig|345219.18.peg.470, fig|345219.18.peg.471
Carbohydrates	Monosaccharides	D-Galacturonate and D-Glucuronate Utilization	D-mannonate oxidoreductase (EC 1.1.1.57)	fig|345219.18.peg.472
Carbohydrates	Monosaccharides	D-Galacturonate and D-Glucuronate Utilization	Mannonate dehydratase (EC 4.2.1.8)	fig|345219.18.peg.473
Carbohydrates	Monosaccharides	D-Tagatose and Galactitol Utilization	PTS system, galactitol-specific IIA component (EC 2.7.1.69)	fig|345219.18.peg.753
Carbohydrates	Monosaccharides	D-Tagatose and Galactitol Utilization	PTS system, galactitol-specific IIB component (EC 2.7.1.69)	fig|345219.18.peg.755
Carbohydrates	Monosaccharides	D-Tagatose and Galactitol Utilization	PTS system, galactitol-specific IIC component (EC 2.7.1.69)	fig|345219.18.peg.754
Carbohydrates	Monosaccharides	D-galactarate, D-glucarate and D-glycerate catabolism	D-glycerate transporter (predicted)	fig|345219.18.peg.901
Carbohydrates	Monosaccharides	Mannose Metabolism	Mannose-1-phosphate guanylyltransferase (GDP) (EC 2.7.7.22)	fig|345219.18.peg.3155
Cell Wall and Capsule	Capsular and extracellular polysacchrides	Exopolysaccharide Biosynthesis	Glycosyl transferase, group 2 family protein	fig|345219.18.peg.3518
Cofactors, Vitamins, Prosthetic Groups, Pigments	Folate and pterines	5-FCL-like protein	YjbQ (alternate ThiE)	fig|345219.18.peg.758
Cofactors, Vitamins, Prosthetic Groups, Pigments	Riboflavin, FMN, FAD	Riboflavin, FMN and FAD metabolism in plants	Molybdopterin binding motif, CinA N-terminal domain	fig|345219.18.peg.1750

5.3.2.1 Starch degradation by Bacillus coagulans ABQ

Following characterisation of the alpha-amylase protein in Chapter 4, *B. coagulans* ABQ was sequenced using the Illumina platform. Starch metabolism in *Bacillus coagulans* ABQ was high at 40 °C. Five enzymes were shown to be present in the breakdown of starch or maltose to three forms of glucose: D-Glucose, α-D-Glucose and α-D-Glucose-1P (Figure 5.7). Alpha-amylases (EC 3.2.1.1), such as the *B. coagulans* ABQ enzyme characterised in Chapter 4, have been shown to degrade starch to dextrin (polymers of D-Glucose) via endo hydrolysis of 1,4 glycosidic linkages (Ballschmiter *et al.*, 2005; R. Gupta *et al.*, 2003). Oligo-1,6-glucosidase (EC 3.2.1.10) breaks down 1,6 glycosidic bonds to reduce remaining dextrins to α-D-Glucose (Watanabe, Hata, Kizaki, Katsube, & Suzuki, 1997). As detailed in Table 5.3, *B. coagulans* ABQ also had an alpha-glucosidase (EC 3.2.1.20), which was required for D-Glucose production. Starch metabolism pathways were observed using RAST Seed Viewer and KAAS (KEGG Automatic Annotation Server).

The Maltose and Maltodextrin Utilization subsystem was selected in *B. coagulans* ABQ genome in the RAST Seed Viewer. Enzymes within this pathway were highlighted in Figure 5.7. All genes within this pathway, including locations on contigs plus up/downstream genes, were detailed in Table 8.6 (Appendix 8.3). *B. coagulans* had an incomplete Maltose and Maltodextrin Utilization pathway. However, the isolate had a complete starch degradation pathway (purple colouring in Figure 5.7) including two cytoplasmic alpha-amylases (EC 3.2.1.1), one of which was characterised in chapter 4 (section 4.2.7). These alpha-amylase genes were then modelled (5.3.5.3).

Genes involved in starch degradation and the maltose and maltodextrin utilisation subsystem in *B. coagulans* were spread over 11 contigs. The cytoplasmic alpha-amylase genes were found on contigs 39 (amylase 2) and 24 (amylase 1) (Appendix 8.3.1). Copies of transporter protein genes (MalE/F/G) were located on contigs 9 and 33. Transporters allow dextrins to pass into the cell and be broken down to α-D-glucose (Figure 5.7).

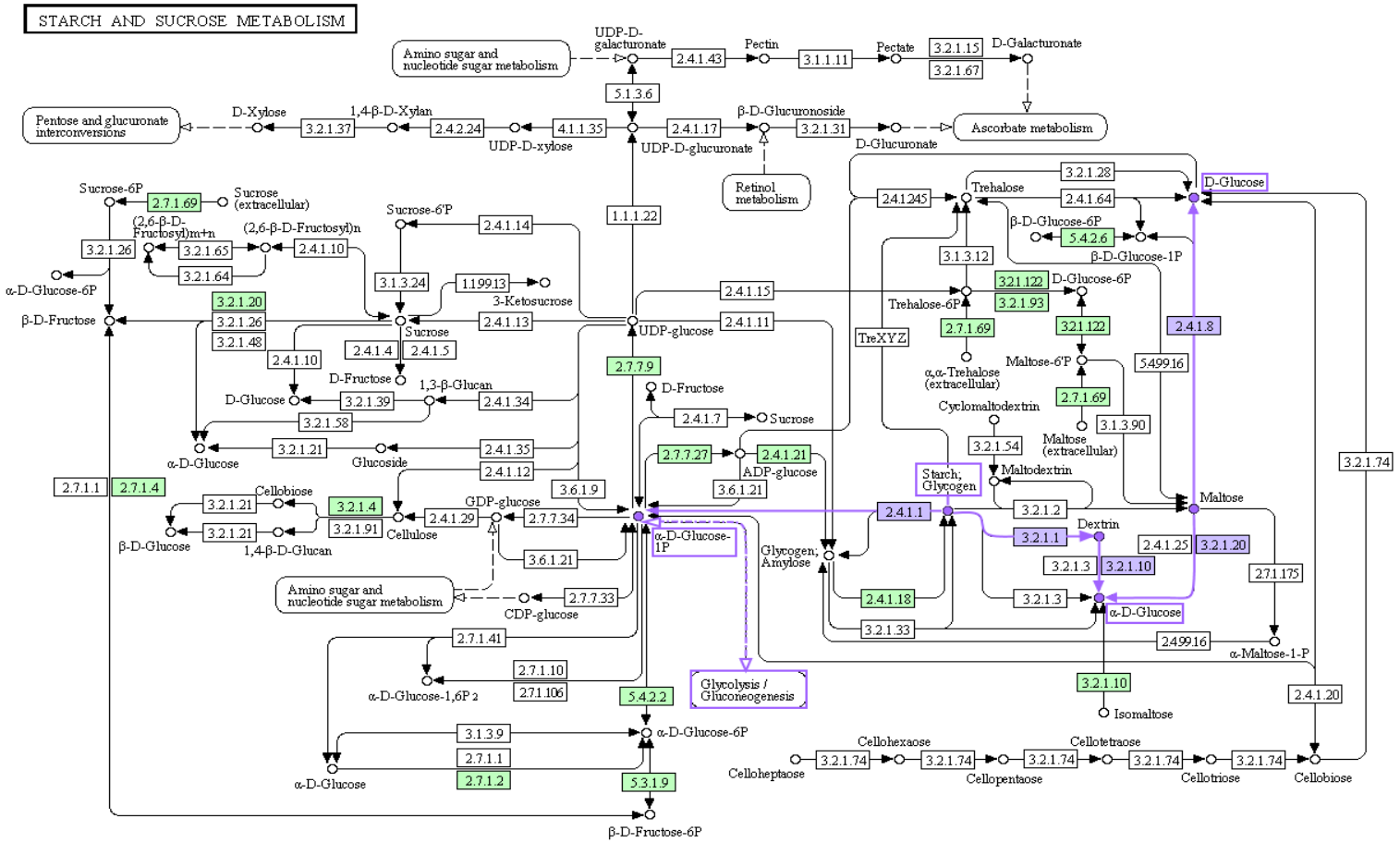


Figure 5.7 Starch and sucrose metabolism in *Bacillus coagulans* ABQ.

Sugar releasing metabolism of starch or glycogen to glucose was highlighted in purple, where as other genes present in *B. coagulans* ABQ were highlighted in green. Generated through KEGG-KAAS via GHOSTZ annotation using contigs converted into amino acids (Kanehisa Laboratories, 2015).

5.3.3 Xylan degradation analysis of Illumina sequenced bacterial isolates

All 23 bacterial isolates were initially chosen due to their ability to break down xylan at pH extremes (section 3.2.5). The xylose utilisation subsystem was compared in RAST Seed Viewer across all isolates. Bacteria isolated at room temperature did not contain genes coding for endo-beta-xylanases (XynA) or beta-xylosides transporter (XynT). Some bacteria isolated at higher temperature also did not contain these genes. However, when XynA and XynT were absent "xylanases" and ABC transporters were present. Enzymes required for xylan degradation to the pentose phosphate pathway are highlighted in Table 5.4, Figure 5.8, and Figure 5.9. XylE/F/G/H/T or XynT/ABC could be used as transporters, and XylS/3/XynA or Xyl could be initial depolymerising enzymes. Some enzymes may have been missed in sequencing.

Interestingly *B. licheniformis* AJF1 and *Brevibacillus brevis* AHW had no enzymes associated with xylose utilisation (Table 5.4). However, they were active on xylan (section 3.2.5.2). Though *Brevibacillus brevis* AHW was only active at pH 11 at 50 °C, *B. licheniformis* AJF1 was active across the pH and temperature range.

Table 5.4 All bacterial isolates from genome sequencing compared in Xylose Utilisation.

Endo-beta-xylanases (XynA) or beta-xylosides transporter (XynT) were not identified in mesophilic isolates, however "xylanases" were identified. *+ – gene present; - – gene not present. * - enzymes required for xylan degraradtion to the pentose phosphate pathway. XylA - Xylose isomerase (EC 5.3.1.5); XylB - Xylulose kinase (EC 2.7.1.17); XylR - Xylose-responsive transcription regulator, ROK family/ Xylose activator XylR (AraC family); XylE - D-xylose proton-symporter; XylF - Xylose ABC transporter, periplasmic xylose-binding protein; XylG - D-xylose transport ATP-binding protein; XylH - Xylose ABC transporter, permease protein; XylT - D-xylose proton-symporter; XynT - Xyloside transporter; ABC - Xylose ABC transporter, substrate-binding component/ Xylose ABC transporter, permease component/ Xylose ABC transporter, ATP-binding component; XylS - Alpha-xylosidase (EC 3.2.1.-); Xyl3 - Beta-xylosidase (EC 3.2.1.37); XynA - Endo-1,4-beta-xylanase A precursor (EC 3.2.1.8); XDH - D-xylose 1-dehydrogenase (EC 1.1.1.175); XL - Xylonolactonase (EC 3.1.1.68); KSAD - Ketoglutarate semialdehyde dehydrogenase (EC 1.2.1.26); Xyl – Xylanase.*

Organism	Active	XylA*	XylB*	XylR	XylE/F/G/H/T*	XynT/ABC*	XylS/3/XynA*	XDH	XL	KSAD	Xyl*
B. coagulans ABQ	+	-	+	-	-	+	-	-	-	-	-
B. licheniformis AHY1	+	+	+	+	+	+	+	-	-	+	-
B. licheniformis AIB	+	+	+	+	-	+	+	-	-	+	-
B. licheniformis AIE1	+	+	+	+	+	+	+	-	-	+	-
B. licheniformis AJF1	-	-	-	-	-	-	-	-	-	-	-
B. licheniformis AJG	+	+	+	+	-	+	+	-	-	+	-
B. licheniformis AJH1	+	+	+	+	+	+	+	-	-	+	-
B. licheniformis AJN1b	+	+	+	+	+	+	+	-	-	+	-
B. licheniformis AJT	+	+	+	+	+	+	+	-	-	+	-
B. licheniformis AJVa	+	+	+	+	+	+	+	-	-	+	-
B. pumilus AIF	+	+	+	+	-	+	+	-	-	-	-
B. pumilus AIK	+	+	+	+	-	+	+	-	-	-	-
B. subtilis AHU	+	+	+	+	-	+	+	-	-	+	-
B. subtilis AJF2	+	+	+	+	-	+	+	-	-	+	-
B. subtilis AJK	+	+	+	+	-	+	+	-	-	+	-
B. subtilis AJN1a	+	+	+	+	-	+	+	-	-	+	-
B. brevis AHW	-	-	-	-	-	-	-	-	-	-	-
B. fungorum AELa	+	+	+	+	+	-	-	-	-	+	-
Burkholderia sp. AEW2a	+	+	+	+	+	-	-	-	-	+	+
S. asaccharolytica AEH1	+	+	+	-	+	+	+	+	+	+	+
L. rhizovicinus AEOb	+	+	+	+	+	-	-	-	-	-	+
L. rhizovicinus AEZa	+	+	+	+	+	-	-	-	-	-	+
L. rhizovicinus AFDa	+	+	+	+	+	-	-	-	-	-	+

5.3.3.1 Xylan metabolism at low pH

Six of the isolates demonstrating xylanase activity at pH 4 were sequenced (Figure 3.8). *S. asaccharolytica* AEH1 had low activity on pH 4 xylan and starch at pH 6.2 only. *L. rhizovicinus* AEOb however, had consistent xylanase activity from pH 4-8.5. Thermotolerant isolates were not active at low pH, but were very active at high pH as well as temperature (section 3.2.5.2). Genes involved in the xylose utilisation pathway were contained to three contigs (Table 8.7). Eight genes were reported to be involved in full xylan degradation to D-xylulose-5P, which leads into the Pentose Phosphate Pathway (Figure 5.8). XynA (endo-1,4-beta-xylanase A precursor (EC 3.2.1.8)) was not identified in *L. rhizovicinus* AEOb. However, two xylanases (Xyl) were present. After BLAST searches with both genes, they were identified as EC 3.2.1.8 and modelled (section 5.3.5.2). As XynT (xyloside transporter) was not present, beta-xylosides could not be transported into the cell with this gene, however xylose transporters (XylF/G/H) were present (Figure 5.8). Therefore, the xylan polymer would most likely be depolymerised to monomeric xylose before transport across the cell wall.

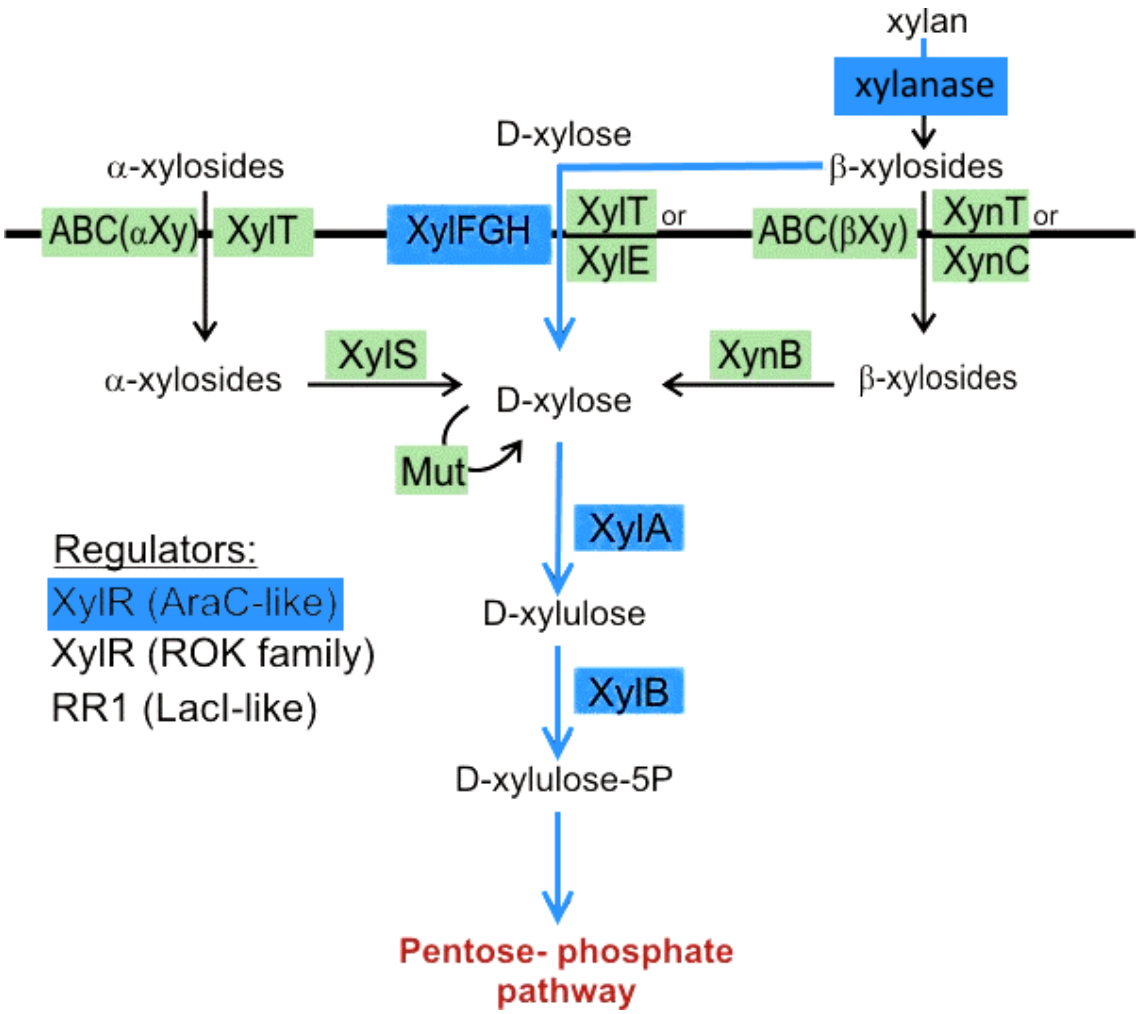


Figure 5.8 Xylose utilisation pathway (Aziz *et al.*, 2008; National Microbial Pathogen Data Resource, 2015; Overbeek *et al.*, 2014).

Enzymes highlighted in blue were present in *L. rhizovicinus* AEOb forming a full pathway of the XylAB pathway (Isomerase pathway). Metabolic products such as ethanol could then be formed through the Pentose phosphate pathway. Green shows other genes present in the xylose utilisation pathway but not present in the AEOb genome.

Xylanases found on contig 42 (fig|6666666.106052.peg.2773: xylanase 1) and 489 (fig|6666666.106052.peg.3024: xylanase 2) of *L. rhizovicinus* AEOb were aligned using NCBI BLASTp to find homologous genes (Table 5.5 and

Table 5.6 respectively). Gene sequences were then aligned to the most homologous hits using ClustalW2 (Appendix 8.3.10 and 8.3.11). Xylanase 1 hit carboxylic esterase, which acts on branches from the xylan backbone (Hesse, Liebert, & Heinze, 2006; Martins *et al.*, 2011). However, other database hits suggested it could actually be an endo-1,4-beta-xylanase (Table 5.5). Xylanase 2 also aligned to endo-1,4-beta-xylanases (Table 5.6).

Table 5.5 Top ten NCBI protein BLAST hits for *L. rhizovicinus* AEOb xylanase 1 found on contig 42 (fig|6666666.106052.peg.2773).

Query coverage was relatively high but identity was low and the most homologous hits were hypothetical proteins from *Luteibacter sp.*.

Description	Max. score	Total score	Query cover.	E value	Ident.	Accession
MULTISPECIES: hypothetical protein [Luteibacter]	414	414	94%	4e-141	72%	WP_036112109.1
hypothetical protein [Luteibacter sp. 9133]	413	413	94%	1e-140	71%	WP_043694746.1
hypothetical protein [Luteibacter yeojuensis]	396	396	90%	7e-134	70%	WP_045829244.1
esterase, partial [Solimonas variicoloris] [Singularimonas variicoloris]	357	357	90%	4e-119	62%	WP_039748441.1
esterase, partial [Solimonas flava] [Sinobacter flavus]	355	355	89%	2e-118	63%	WP_043112903.1
esterase [Hydrocarboniphaga effusa]	337	337	89%	2e-111	59%	WP_040454937.1
hypothetical protein WQQ_35120 [Hydrocarboniphaga effusa AP103]	338	338	90%	4e-111	58%	EIT68317.1
Endo-1,4-beta-xylanase B B [Stenotrophomonas maltophilia SKK35]	324	324	84%	1e-105	60%	CCP11714.1
Esterase/lipase [Stenotrophomonas maltophilia EPM1]	321	321	84%	2e-104	60%	EMF60636.1
esterase [Stenotrophomonas maltophilia]	319	319	83%	3e-104	60%	WP_032967165.1

Table 5.6 Top ten NCBI protein BLAST hits for *L. rhizovicinus* AEOb xylanase 2 found on contig 489 (fig|6666666.106052.peg.3024).

Query coverage was relatively high but identity was low and the most homologous hits were from a range of bacteria.

Description	Max. score	Total score	Query cover.	E value	Ident.	Accession
xylanase [Caulobacter sp. JGI 0001010-J14]	324	324	96%	1e-106	60%	WP_029589220.1
hypothetical protein [Acidobacteriaceae bacterium KBS 96]	321	321	94%	6e-106	61%	WP_020720866.1
xylanase [Stenotrophomonas maltophilia JV3]	321	321	94%	1e-105	60%	AEM52257.1
xylanase [Stenotrophomonas maltophilia]	321	321	94%	1e-105	60%	WP_042359270.1
xylanase [Stenotrophomonas maltophilia]	318	318	94%	2e-104	60%	WP_032961959.1
xylanase [Stenotrophomonas maltophilia]	317	317	94%	4e-104	60%	WP_033831700.1
xylanase [Caulobacter henricii]	316	316	98%	8e-104	57%	WP_035088150.1
xylanase [Caulobacter henricii]	316	316	98%	1e-103	57%	WP_035080244.1
xylanase [Stenotrophomonas sp. RIT309]	316	316	94%	1e-103	60%	WP_032975056.1
xylanase [Stenotrophomonas maltophilia]	314	314	99%	5e-103	57%	WP_006382208.1

5.3.3.2 Xylan degradation at high pH

Bacillus licheniformis AHY1 was chosen for whole genome sequencing due to its high activity on xylan agar across the pH range. There was activity at pH 6.2 and 8.5 at 40 °C; high activity across all pHs (bar pH 4) at 50 °C. Low activity was present at pH 6.2 but high at 8.5 at 60 °C, and low activity at pH 6.2 and 8.5 at 70 °C (section 4.1.3).

Genes associated with xylan degradation were found on six contigs (Appendix 8.3.3.2). The pathway used in *B. licheniformis* AHY1 for xylan degradation within the xylose utilisation subsystem is highlighted in Figure 5.9. Endo-1,4-beta-xylanase (XynA) initiates breakdown of xylan in most systems. However, no endo-1,4-beta-xylanase (EC 3.2.1.8) gene was identified in *B. licheniformis* AHY1. RAST annotations indicated that

six out of the eight *B. licheniformis* isolates uploaded did not have this gene. *B licheniformis* ATTC 14580 also did not contain this gene. Both *B. licheniformis* AJG and AIB had genes encoding an enzyme that RAST classed as endo-1,4-beta-xylanase. After a blastp search these hit arabinoxylan arabinofuranohydrolase in *Bacillus sp.* (AIB arabinoxylan arabinofuranohydrolase: 100% query coverage, 0.0 e value and 99 % identity to *B. subtilis*; AJG (XynD): 100% query coverage, 0.0 e value and 100% max identity to *B. lichenifomis*). *B. licheniformis* AJG also had another gene classed as endo-1,4-beta-xylanase which aligned with hypothetical proteins but also hit xylanase Y in *B. licheniformis* (WP_003180630.1). UniProt classified XylY (Glycoside Hydrolase Family 8) as having endo-1,4-beta-xylanase activity in *Bacillus sp.* (Appendix 8.3.12). However, two beta-xylosidases (Xyl3) were identified in *B. licheniformis* AHY1. Both genes had exact homology to glycoside hydrolases (Appendices 8.3.2). fig6666666106061peg3747 (beta-xylosidase 1) had exact homology to Q65MB7 (Glycoside Hydrolase Family 43), and fig6666666106061peg3748 (beta-xylosidase 2) had exact homology to Q65MB6 (xynB, Glycoside Hydrolase Family 43).

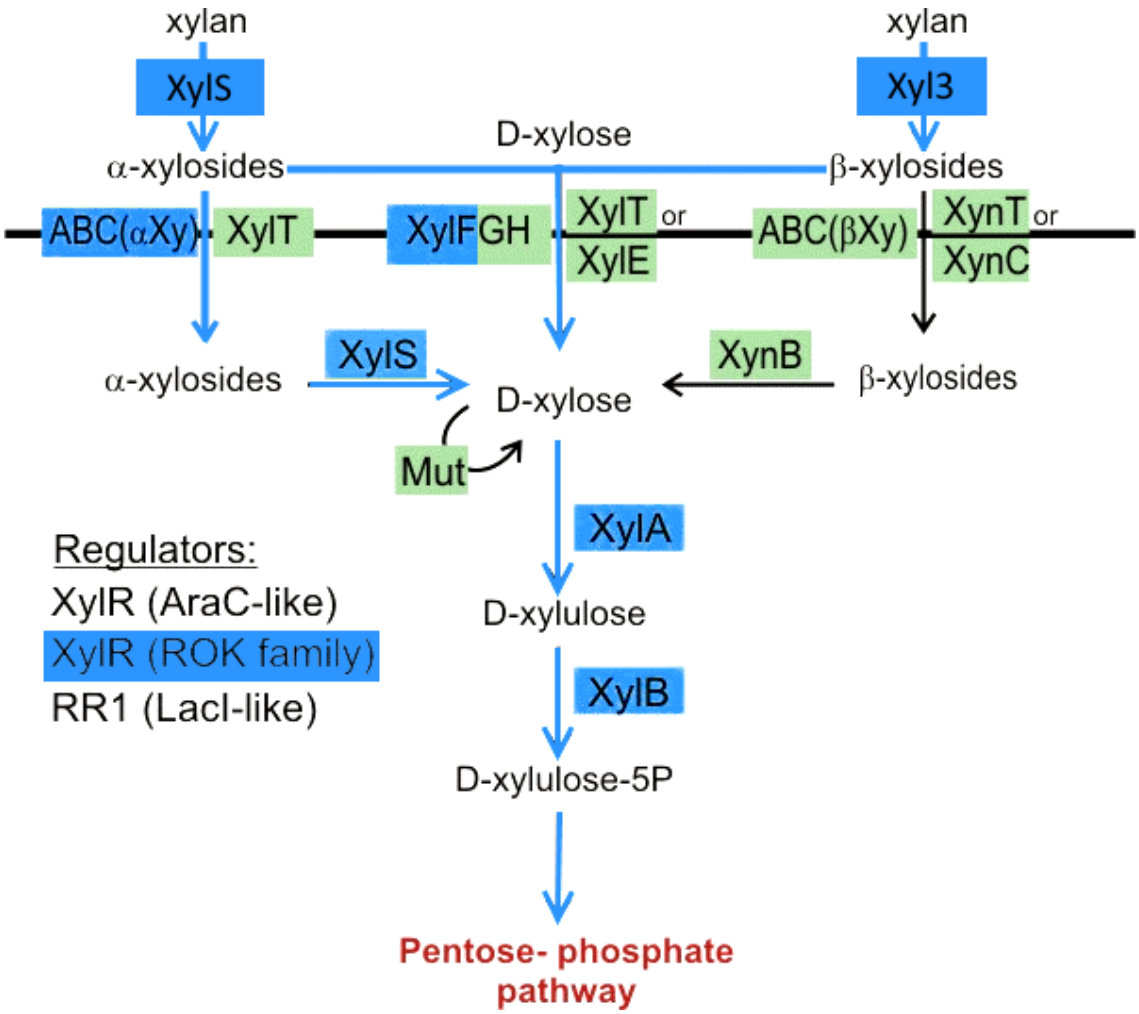


Figure 5.9 Xylose utilisation pathway (Aziz *et al.*, 2008; National Microbial Pathogen Data Resource, 2015; Overbeek *et al.*, 2014) in *Bacillus licheniformis* AHY1.

XynA was not present in the AHY1 genome. However Xyl3 (Beta-xylosidase (EC 3.2.1.37)) and XylS (alpha-xylosidase (EC 3.2.1.-)) degrade the xylan backbone to alpha- and beta-xylosides. Alpha-xylosides and xylose could be transported into the cell via ABC transporters (ABCa and XylF) and further degraded. Enzymes highlighted in blue were present in AHY1. Green shows other genes present in the xylose utilisation pathway but not present in the AHY1 genome.

5.3.4 Cellulose metabolism by *Bacillus subtilis* AJF2

Bacillus subtilis AJF2 had high activity on CMC agar at 40-50 °C at pH 6.2-10. There was still activity at 60-70 °C at pH 6.2-8. A number of subsystem pathways were viewed using RAST Seed Viewer. Finally the Beta-Glucoside Metabolism pathway was used as AJF2 had the Cel(cellobiose) operon and was most probably used in cellulose breakdown. AJF2 has a full pathway using this operon which was spread over 7 contigs (Appendix 8.3.9).

CelA (Cytyoplasmic endo-1,4-beta-glucanase (EC 3.2.1.4)) was not found in RAST annotations. However, a endo-1,4-beta-glucanase was highlighted when viewing pathways in KAAS (fig|6666666.106072.peg.3365, contig 6, 1'503 bp: cellulase 2). The only glucanase found using RAST annotations was EC 3.2.1.73 (fig|6666666.106072.peg.2457, licheninase, GH16 superfamily), which was highly homologous to *B. subtilis* BSn5 (one difference in amino acid sequence, two differences in nucleotide sequence). BLASTp of the *B. subtilis* AJF2 lichenninase gene had high homology to beta-glucanase and a cellulase from *B. subtilis* (Table 5.7). Top hits were aligned with *B. subtilis* AJF2 and BSn5 EC 3.2.1.73 genes resulting in high homology (Appendix 8.3.6).

Using KAAS, gene sequences cellulase 2 and fig|6666666.106072.peg.1834 (cellulase 1) were identified as beta-glucanase (EC 3.2.1.4). However, BLASTp of *B. subtilis* H6WUK0 cellulase (EC 3.2.1.4) against AJF2 contigs was highly homologous to cellulase 2. When aligned with *B. subtilis* MB73/2 M2UBG4 cellulase (EC 3.2.1.4) *B. subtilis* AJF2 cellulase 2 had only two different amino acids (Appendix 8.3.6). BLASTp of *B. subtilis* contigs against a number of beta-glucanase genes revealed exact homology with *B.*

subtilis QB928 J7JXX5 cellulase (EC 3.2.1.4) but very little homology with cellulase 2 (Appendix 8.3.7).

Table 5.7 Top seven BLASTp results for *Bacillus subtilis* AJF2 licheninase (EC 3.2.1.73) search using NCBI.

Query coverage and identity had 100% homology to beta-glucanases from *B. subtilis*, and the e value was very low.

Description	Max. score	Total score	Query cover	E value	Ident.	Accession
beta-glucanase [Bacillus subtilis]	503	503	100%	7e-179	100%	WP_024571825.1
beta-glucanase [Bacillus subtilis]	502	502	100%	2e-178	99%	WP_003244531.1
beta-glucanase [Bacillus subtilis]	502	502	100%	2e-178	99%	WP_015483903.1
cellulase [Bacillus subtilis]	501	501	100%	3e-178	99%	AHZ57098.1
beta-glucanase [Bacillus subtilis]	501	501	100%	6e-178	99%	WP_038828911.1
beta-glucanase [Bacillus sp. A053]	501	501	100%	7e-178	99%	WP_040081344.1
beta-glucanase [Bacillus subtilis]	501	501	100%	7e-178	99%	WP_014478425.1

Lichenin is made up of repeating glucose units linked by β-1,3 and β-1,4 glycosidic bonds (Perlin & Suzuki, 1962). Beta-glucoside consists of two glucose units bound by beta-1,4 links. Cellulose degradation most likely occurs using beta-glucosidase (Lymar, Li, & Renganathan, 1995) as highlighted in Figure 5.10. *Bacillus subtilis* AJF2 had two copies of beta-glucosidase (EC 3.2.1.21) (LicH) and two copies of 6-phospho-beta-glucosidase (EC 3.2.1.86) (BglB). AJF2 most likely utilised the LicBCAH operon with a full suite of protein encoding genes. Utilisation of beta-glucosidic compounds uses five genes: phosphotransferase system (PTS) enzyme II (licB and licC) and enzyme IIA (licA), a potential 6-phospho-beta-glucosidase (licH), as well as a putative regulator protein (licR) (Tobisch *et al.*, 1997). AJF2 had all five genes across four contigs (Appendix 8.3.3).

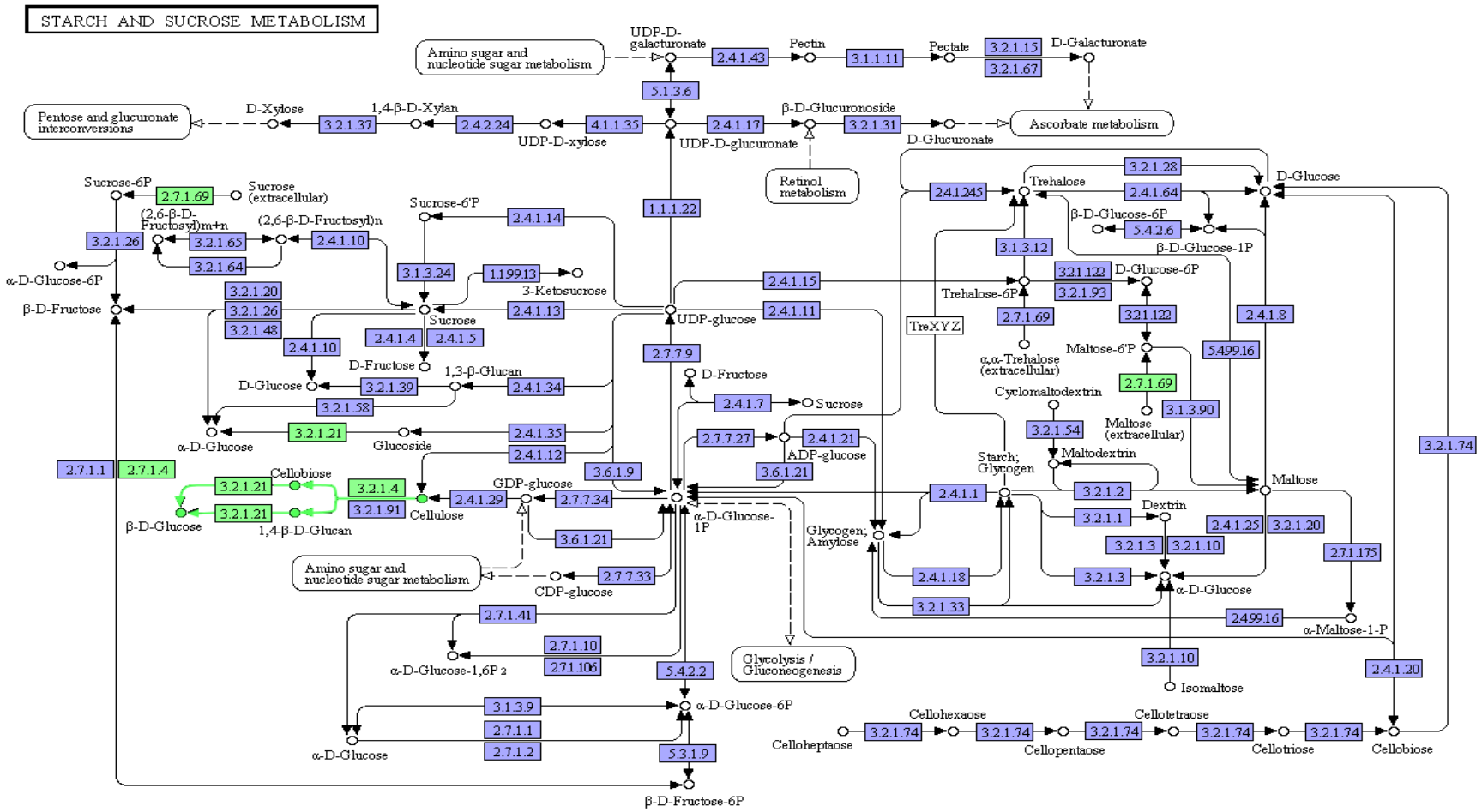


Figure 5.10 Cellulose degradation by *Bacillus subtilis* AJF2.

Enzymes in the starch and sucrose metabolism pathway present in *B. subtilis* AJF2 are highlighted in green, green arrows show the cellulose degradation pathway. Pathway drawn by KAAS.

5.3.5 Protein models of nine genes involved in monomeric sugar release from cellulose, xylan and starch

Nine protein sequences, identified as sugar degrading enzyme genes by RAST, were modelled through Phyre2 software (Kelley & Sternberg, 2009). Pdb files were submitted to 3DLigandSite for binding site predictions (Wass *et al.*, 2010). Modelled proteins are listed in Table 5.8, whereby they were submitted to NCBI CDD for confirmation of identification and functionality. Four proteins were predicted to be incomplete at either the N terminus, C terminus or both termini. Though incompleteness was only predicted at multi-domain or superfamily level.

A summary of protein models can be seen in Table 5.9, based on results from three databases: Phyre2, 3DLigandSite and SignalP. Though confidence is high for each model, coverage compared to the reference model ranges from 54%-99%. However, only proteins with crystal structures (in PDB) were available as reference models meaning models were hypothetical and binding sites were not always accurate.

Table 5.8 NCBI Conserved Domain Database (CDD) results for nine enzymes chosen for protein modelling.

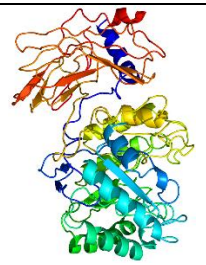
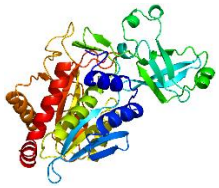
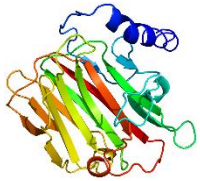
PSSM – position specific scoring matrix; "-" – complete; N – incomplete at N-terminus; C – incomplete at C-terminus; NC – incomplete at both N- and C-terminus (RPS-BLAST omitted more than 40% of the conserved domains extent).

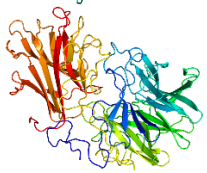
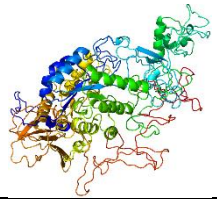
Query	Hit type	PSSM-ID	From	To	E-Value	Bitscore	Accession	Short name	Incomplete	Superfamily
fig|6666666.106072.peg.2457 (*B. subtilis* AJF2 licheninase)	Specific	185684	34	241	4.10529e-122	348.877	cd02175	GH16_lichenase	-	cl00218
	Superfamily	260271	34	241	4.10529e-122	348.877	cl00218	Glyco_hydrolase_16 superfamily	-	-
	Multi-dom	225182	1	241	4.89257e-69	218.492	COG2273	SKN1	C	-
fig|6666666.106072.peg.1834 (*B. subtilis* AJF2 cellulase 1)	Specific	193532	10	347	0	531.638	cd05656	M42_Frv	-	cl14876
	Superfamily	276098	10	347	0	531.638	cl14876	Zinc_peptidase_like superfamily	-	-
fig|6666666.106072.peg.3365 (*B. subtilis* AJF2 cellulase 2)	Specific	249631	50	296	3.09908e-78	248.029	pfam00150	Cellulase	-	cl07893
	Superfamily	275846	50	296	3.09908e-78	248.029	cl07893	AmyAc_family superfamily	-	-
	Specific	250246	356	437	8.2225e-29	108.93	pfam00942	CBM_3	-	cl03026
	Superfamily	261553	356	437	8.2225e-29	108.93	cl03026	CBM_3 superfamily	-	-
	Multi-dom	225344	9	445	8.33414e-60	203.871	COG2730	BglC	-	-
fig|345219.18.peg.857 (*B. coagulans* ABQ alpha-amylase 2)	Superfamily	275846	41	626	2.72128e-40	153.441	cl07893	AmyAc_family superfamily	-	-
	Multi-dom	236518	41	715	6.75779e-58	205.893	PRK09441	PRK09441	-	-
fig|345219.18.peg.2080 (*B. coagulans* ABQ alpha-amylase 1)	Specific	200457	4	392	0	701.58	cd11318	AmyAc_bac_fung_AmyA	-	cl07893
	Superfamily	275846	4	392	0	701.58	cl07893	AmyAc_family superfamily	-	-
	Superfamily	275776	426	480	0.00146825	36.6316	cl02706	Alpha-amylase_C superfamily	-	-
	Multi-dom	236518	2	481	0	809.886	PRK09441	PRK09441	-	-
fig|6666666.106061.peg.3747 (*B. licheniformis* AHY1 beta-xylosidase 1)	Specific	185742	25	297	1.90461e-128	378.455	cd09001	GH43_XYL_2	-	cl14647
	Superfamily	265439	25	297	1.90461e-128	378.455	cl14647	GH43_62_32_68 superfamily	-	-
	Multi-dom	226038	23	511	5.3232e-74	246.585	COG3507	XynB	-	-

fig|6666666.106061.peg.3748 (*B. licheniformis* AHY1 beta-xylosidase 2)	Specific	185741	9	292	2.01894e-146	424.619	cd09000	GH43_XYL_1	-	cl14647
	Superfamily	265439	9	292	2.01894e-146	424.619	cl14647	GH43_62_32_68 superfamily	-	-
	Superfamily	242478	376	463	0.00629284	36.4531	cl01397	DUF1349 superfamily	NC	-
	Multi-dom	226038	6	512	2.39971e-120	367.153	COG3507	XynB	-	-
fig|6666666.106052.2773 (*Unknown sp.* AEOb xylanase 1)	Superfamily	276326	75	279	8.36885e-10	57.8062	cl21494	Esterase_lipase superfamily	N	-
	Multi-dom	223730	72	298	4.12259e-26	105.017	COG0657	Aes	N	-
fig|6666666.106052.3024 (*Unknown sp.* AEOb xylanase 2)	Superfamily	276326	114	231	2.98664e-06	45.6715	cl21494	Esterase_lipase superfamily	N	-
	Superfamily	276326	20	133	0.00221112	37.6968	cl21494	Esterase_lipase superfamily	C	-
	Multi-dom	257229	34	229	5.0994e-07	46.5564	pfam12695	Abhydrolase_5	-	-

Table 5.9 Modelled proteins summary.

Using Phyre2 (top model, confidence, coverage, template information, transmembrane helix), 3DLigandSite (binding site residues) and SignalP 3.0 (signal peptide). Models coloured by rainbow N-C terminus.

Bacterial isolate	Name of enzyme	RAST /Phyre2 protein name	Top model	Confidence (%)	Coverage (%)	Top template information	Binding site residues	Trans-membrane helix	Signal peptide	Phyre2 Intensive model
B. subtilis AJF2	Cellulase (EC 3.2.1.4)	fig|6666666.106072.peg.3365/f695ba4f52eeb880	d1g01a	100	63	**Family**: Beta-glycanases **Superfamily:** (Trans)glycosidases **Fold:** TIM beta/alpha-barrel **Chain:** A **Molecule:** endoglucanase	65 (His), 69 (Trp), 70 (Trp), 96 (Tyr), 99 (Asp), 133 (Leu), 168 (Asn), 169 (Glu), 257 (Glu), 291 (Trp), 296 (Lys)	Yes	Yes	
B. subtilis AJF2	Cellulase (EC 3.2.1.4)	fig|6666666.106072.peg.1834/8819d03f6bdd5d30	c1vheA	100	99	**Chain:** A **Molecule:** aminopeptidase/glucanase	68 (His), 182 (Asp), 214 (Glu), 215 (Glu), 237 (Asp), 325 (His)	No	No	
B. subtilis AJF2	Endo-beta-1,3-1,4 glucanase (Licheninase) (EC 3.2.1.73)	fig|6666666.106072.peg.2457/011e26063a3949f0	d1gbga	100	88	**Family:** Glycosyl hydrolase family 16 **Superfamily:** Concanavalin A-like lectins/glucanases **Fold:** Concanavalin A-like lectins/glucanases **Chain:** A	37 (Pro), 73 (Gly), 235 (Asp)	Yes	Yes	

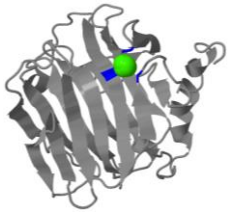
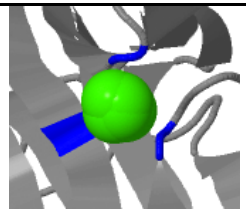
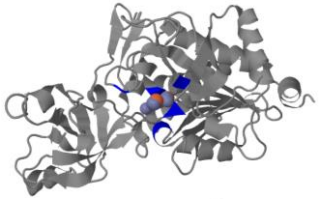
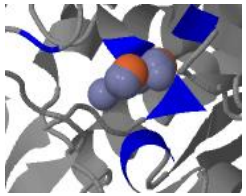
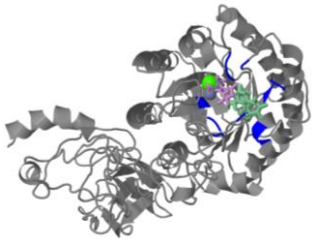
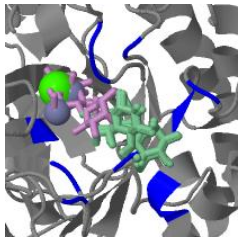
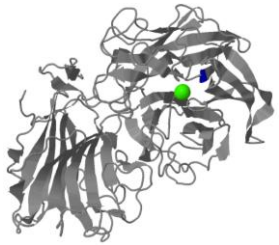
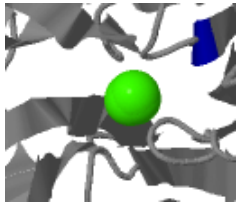
						Molecule: (1,3-1,4)-beta-d-glucan 4 glucanohydrolase				
Unknown sp. AEOb	Xylanase (EC 3.2.1.8)	fig|6666666.106052.peg.3024/875e3973775aa2c5	c3bxpA	100	95	**Chain:** A **Molecule:** putative lipase/esterase	85 (Asp), 88 (Arg)	No	No	
Unknown sp. AEOb	Xylanase (EC 3.2.1.8)	fig|6666666.106052.peg.2773/e9a3327569e4f970	c3bxpA	100	78	**Chain:** A **Molecule:** putative lipase/esterase	*Insufficient homologues*	Yes	Yes	
Bacillus licheniformis AHY1	Beta-xylosidase (EC 3.2.1.37)	fig|6666666.106061.peg.3747/678778c578bb2a06	c1yifC	100	94	**Chain:** C **Molecule:** beta-1,4-xylosidase	204 (Lys)	No	No	
Bacillus licheniformis AHY1	Beta-xylosidase (EC 3.2.1.37)	fig|6666666.106061.peg.3748/9f524431b2263ea2	c1yifC	100	98	**Chain:** C **Molecule:** beta-1,4-xylosidase	347 (Gly), 508 (Asp)	No	No	
Bacillus coagulans ABQ	Cytoplasmic alpha-amylase (EC 3.2.1.1)	fig|345219.18.peg.2080/73510eb1372cfbf1	c1hvxA	100	98	**Chain:** A **Molecule:** alpha-amylase	13 (Trp), 56 (Tyr), 199 (Met), 200 (Tyr), 330 (Asp),	Yes	No	
Bacillus coagulans ABQ	Cytoplasmic alpha-amylase (EC 3.2.1.1)	fig|345219.18.peg.857/312e3c1131f89c2f	c3aicC	100	54	**Chain:** C **Molecule:** glucosyltransferase-si	353 (Leu), 354 (Phe), 382 (Asp)	Yes	Yes	

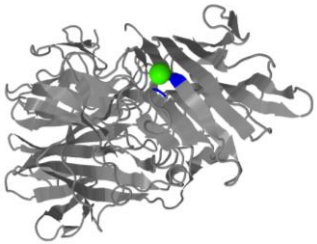
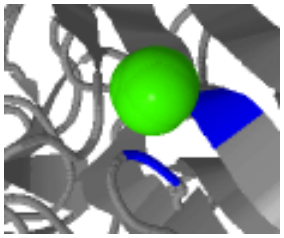
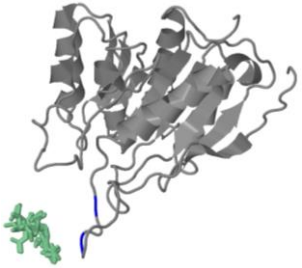
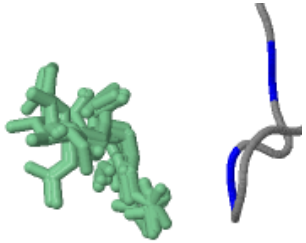
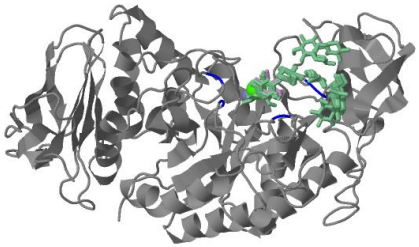
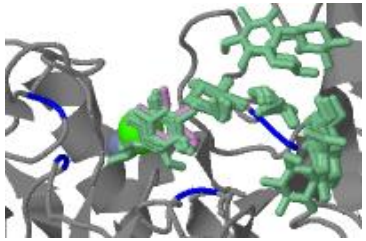
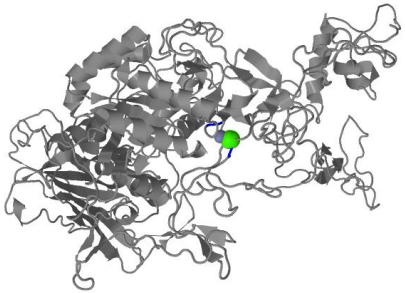
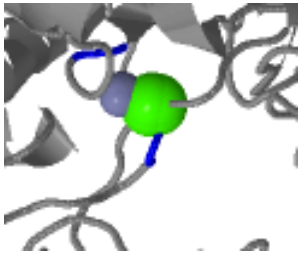
5.3.5.1 Cellulose degrading enzymes found in Bacillus subtilis AJF2

The licheninase (EC 3.2.1.73) identified by RAST in the *B. subtilis* AJF2 genome was modelled using Phyre2 Intensive (Table 5.9). Two top models found were d1gbga glycosyl hydrolase family 16 and c3o5sA endo-beta-1,3-1,4-glucanase from *B. subtilis subsp. subtilis* 168. A PSI (Position-Specific Iterated)-Blast Pseudo-multiple sequence alignment was also run for sequence homology. The top hit (UniRef50_A3DJ79, e-value 1e-49) was aligned to the licheninase amino acid sequence and the binding site highlighted in red (8.3.9 Appendix 8). The Phyre2 model was then submitted to 3DLigandSite for ligand binding site prediction (Table 5.9 and Table 5.10). The licheninase was modelled using Phyre2 Intensive (Kelley & Sternberg, 2009). Two cellulase gene sequences were also identified in the *B. subtilis* AJF2 genome by RAST. Both amino acid sequences were aligned with a known cellulase in Appendix 8.3.9. Similarities could be seen between the sequences though the predicted binding sites do not occur in the conserved regions. Both genes were also modelled using Phyre2 (Table 5.9) then submitted to 3DLigandSite for binding site prediction (Table 5.10).

Table 5.10 Binding site models for Phyre2 Intensive modelled proteins.

3DLigandSite highlighted amino acid residues involved in binding (blue). *ND – not determined.*

Enzyme	3DLigandSite model	Binding site	Metallic heterogens	Non-metalic heterogens	Dimensions
Licheninase			Calcium	ND	X:52.439, Y:46.688, Z:51.540
Cellulase 1			Iron, zinc	ND	X:50.393, Y:49.080, Z:72.115
Cellulase 2			Calcium, zinc	beta-D-mannose, beta-D-galactose, alpha-D-mannose, alpha-D-glucose	X:78.494, Y:54.705, Z:56.284
Beta-xylosidase 1			Magnesium, calcium		X:81.581, Y:53.767, Z:61.991

Beta-xylosidase 2			Calcium	ND	X:57.396 Y:57.734 Z:73.391
Xylanase 2			ND	N-acetyl-D-glucosamine	X:53.780, Y:46.531, Z:47.530
Alpha-amylase 1			Calcium, zinc	alpha-D-glucose, configuration of the bound glucose unit	X:83.743, Y:98.382, Z:86.351
Alpha-amylase 2			Calcium, zinc	ND	X:66.374, Y:55.588, Z:74.999

5.3.5.2 Xylan degrading enzymes found in B. licheniformis AHY1 and L. rhizovicinus AEOb

Xylanases could not be found in *B. licheniformis* AHY1; however two beta-xylosidases were present. Both were modelled using Phyre2 Intensive for a 3D protein model. Models were then submitted to 3DLigandSite for binding site analysis. Both protein models for beta-xylosidases had few residues that made up the predicted ligand binding site. Beta-xylosidase 1 had only a single residue identified and beta-xylosidase 2 was shown to have only two residues (Table 5.9). Protein sequences for both xylosidases were aligned with top hits from PSI-Blast pseudo-multiple sequence alignment and binding site residues highlighted in red (Appendix 8.3.12).

Xylanase 2 from *L. rhizovicinus* AEOb (Table 5.9 and Table 5.10) had a top model of c3bxpA putative lipase/esterase and top PSI-Blast hit of UniRef50_B3VP83 (e-value 1e-55) from *Lactobacillus plantarum* wcfs1. The protein sequence was aligned with the top hit, and ligand binding site identified by 3DLigandSite and highlighted in red in the animo acid sequence alignment (Appendix 8.3.11). Binding site predictions could not be determined by 3DLigandSite for *L. rhizovicinus* AEOb xylanase 1. There were insufficient homologous structures with ligands bound found.

5.3.5.3 Starch degrading enzymes found in Bacillus coagulans ABQ

Two cytoplasmic alpha-amylase genes were identified within the *B. cogulans* ABQ genome. The amylase 1 top model was identified as c1hvxA and top PSI-Blast hit was UniRef50_B9E9W0 (e-value 3e-73). The amylase 1 amino acid sequence was aligned with B9E9W0 and the binding site identified by 3DLigandSite (Table 5.10) highlighted in red in Appendix 8.3.13. The top model identified by Phyre2 for cytoplasmic alpha-amylase 2 in ABQ was c3aicC. The best sequence match described by PSI-Blast,

B1YMN6 (e-value 1e-75), was aligned to the gene sequence and predicted binding site highlighted in red (Appendix 8.3.13) as identified by 3DLigandSite (Table 5.10).

5.3.6 Potential acid -tolerant genes

Burkholderia fungorum AELa demonstrated xylanase and amylase activity at pH 4. After functional comparison with *B. fungorum* NBRC 102489 a number of genes were present in AELa genome but not NBRC 102489. D-alanine-poly(phosphoribitol) ligase subunit 1 (EC 6.1.1.13) was one such gene. EC 6.1.1.13 is involved in biosynthesis of D-alanyl-lipoteichoic acid (LTA) and wall teichoic acid (WTA) in Gram-positive bacteria (Perego *et al.*, 1995). *Burholderia fungorum* was known to be Gram-negative, however AELa stained positively. LTA and WTA have been reported to give the cell wall structure, resist attack, assist cell division, transport nutrients, aid bacterial attachment to substrates and participate in cell to cell signalling (Garimella, Halye, Harrison, Klebba, & Rice, 2009). Presence of EC 6.1.1.13 potentially pointed towards a reason for AELa's ability to survive and have active enzymes at low pHs.

5.4 Discussion

5.4.1 Sequencing and analysis of bacterial genomes

Bacterial genomes were sequenced using Illumina MiSeq (San Diego, cat no.: SY-410-1003). Nucleotide sequences were imported to CLC Genomics Workbench v7.1. Sequences were then merged, trimmed and assembled into contigs via de novo assembly. Sequencing depth across the 23 genomes was varied. Lower coverage of genome sequences was observed for *B. licheniformis* AHY1, *B. subtilis* AJF2 and *L. rhizovicinus* AEOb. Of the five previously *Unknown sp.*, one was reclassified as a *B. licheniformis* (Figure 5.6). The remaining four were compared to closest neighbours suggested by RAST but had low homology. However, when compared to AEOb as a

reference genome, AEZa and AFDa had very high homology to AEOb, though AEH1 did not (Figure 5.4). However, BLAST analysis of the 16S rRNA gene found within each genome revealed that AEOb, AEZa and AFDa were most likely to be *L. rhizovicinus* (Figure 5.5). AEH1 was found to be a *Sphingomonas asaccharolytica* strain using the same method. AEOb and AFDa were thought to be mixed samples, however only a single version of the 16S rRNA gene was found within each genome. This could indicate that these bacteria were not mixed samples, were the same species mixed, or sequencing coverage was low. Though, the three genomes were very similar in size of 4.7-4.8 Mbp (Table 5.1) and therefore not likely to be mixed.

Genomes were annotated using RAST. Subsystems of each genome were identified by the Seed Viewer (Appendix 3: 8.3.1). However, annotations were below 50% of the total coverage (Table 5.2). The carbohydrate subsystem was the largest subsystem for each genome bar *Unknown sp.* AEOb, *Burkholderia sp.* AEW2a, *Unknown sp.* AFDa, *Brevibacillus brevis* AHW and *Bacillus licheniformis* AJF1. Here, Amino Acids and Derivatives subsystem was marginally bigger. Whole subsystems could then be viewed and individual genes picked out and modelled. When a single gene was selected its position on the contig, as well as surrounding genes, were detailed.

For economically viable products, such as biobutanol, utilisation of lignocellulose as a fermentation substrate is key. Previous studies have indicated that utilisation of the available xylose has been slow from lignocellulosic hyrolysates (L. Yu *et al.*, 2015). Thermotolerance for enzymes and bacteria is useful for saccharification and fermentation due to the improved solubility and reduced viscosity of sugars (Adams & Kelly, 1998). Cost of fungal enzymes for fermentation can be lowered by using

bacterial counterparts with similar optimum conditions (Rhee *et al.*, 2011). Industrially enzymes, such as cellulases and amylases are used to degrade polymers in detergents as some have been shown to be stable at high pH (Rothschild & Mancinelli, 2001). Also, pre-treatments of biomass are most effective with alkali or acids (section 1.2.3.2), thus wastes and cost could be reduced by saccharification occurring at high or low pH. Therefore, bacterial isolates were selected and genomic DNA sequenced due to their xylanase and amylase activity at extremes of pH and temperature (Figure 3.10 and Figure 3.13-3.28).

Six of the *B. licheniformis* genomes were missing endo-xylanase genes. However, the reference also did not have this gene so an alternative pathway was most likely used. Beta-xylosidases have been shown to be one of three major hemicellulose degrading enzymes that degrade the xylan backbone (H. Cheng & Wang, 2013; Sukumaran, 2009, p. 144). As this gene was identified in *B. licheniformis* AHY1 it was presumed to take the role of primary xylan degrading enzyme. Additionally, genes encoding XynA and XynT were missing from bacteria isolated at room temperature (section 5.3.3). However, ABC transporters and xylanases were present within the genomes. ABC transporters are ATP-binding cassettes and have been shown to bind specifically to xylo-oligosaccharides in thermophilic bacteria (Y. Han *et al.*, 2012; VanFossen, Verhaart, Kengen, & Kelly, 2009). However, they are an important class of transporters in all domains of life (Albers, Koning, Konings, & Driessen, 2004; Koning, Albers, Konings, & Driessen, 2002) including mesophilic bacteria such as *Clostridium cellulolyticum* (de Lucena *et al.*, 2012; Linke *et al.*, 2013; Qian *et al.*, 2003; C. Xu *et al.*, 2013; Z. Zhang, Song, Jiang, Xue, & Ma, 2015). Though, *Clostridium acetobutylicum* and

B. subtilis have been shown to encode only XynT as a transport system, not ABC. It was possible that genes were missed in sequencing, however few bacteria appear to have both XynT and ABC transporter systems (Gu *et al.*, 2010). Of the six mesophilic isolates that were sequenced, *S. asaccharolytica* genome had not been published therefore it was unknown if the absence of XynA and XynT was unusual. However, the *L. rhizovicinus* strain that was published also did not code for either of these gene. Two mesophilic isolates were classified as *Burkolderia sp.,* and based on RAST analysis, no *Burkholderia sp.* were recorded to have genes encoding XynA and XynT.

B. coagulans ABQ had very high coverage; total sequences were 599x the genome size. Therefore, it was unlikely that many genes were missed when sequencing. Also, this strain was highly similar to the reference genome (*B. coagulans* 36D1) and therefore accurately annotated. However, 1515 hypothetical genes were still present in the genome, with only 45% of genes covered in subsystems.

Burkholderia sp. AEW2a had over 13 Mbp of sequenced genome. When aligned to closest neighbour *Burkholderia phytofirmans* PsJN. Homology was spread between regions of low homology (10%) to small highly homologous regions (100%). Bacteria have an average genome size of 4 Mbp (0.6-10 Mbp) (Bentley & Parkhill, 2004; Lewin, 2007; Peterson & Fraser, 2001). Due to this large size there were most likely two or more genomes. The reference *Burkholderia fungorum* NBRC 102489 had a genome size of 8 696 214 bp. Two copies of 16S rRNA gene were present; one was identified as *B. fungorum* (query coverage: 91%, evalue: 0.0, maxiumum identity: 100%) whereas the other was identified as *Luteibacter jiangsuensis* (query coverage: 90%, evalue: 0.0, maxiumum identity: 99%). However, the only non-*Burkholderia sp*. strain listed in

"closest neighbour" was *Desulfatibacillum alkenivorans* AK-01. *L. jiangsuensis* does not have a genome published, therefore AEW2a was aligned to *L. rhizovicinus* DSM 16549, whereby there was low homology.

Workflow was not available in the CLC package when using Blast2Go; therefore there were problems with annotation. "Mapping" did not seem to work until annotation was selected. Also the InterPro scan identified genes for less than half of the genomes contigs. If a general field was applied rather than bacterial genetic code (11), fewer proteins were identified. Contigs were uploaded to RAST for further annotation. As shown in Table 5.2, annotations from RAST Seed Viewer did not cover the whole genome. However, gaps in homology (Figure 5.1 and Figure 5.2) could have been due to miscalling genes or low level of annotations from reference strains. Many genes were not included in subsystems with over 1000 hypothetical proteins per genome, plus hypothetical calls within subsystems. Contigs were ordered and aligned to the reference by RAST Seed Viewer; however this was not entirely accurate. A better way to achieve assembly would be to use Mauve, whereby whole genomes could be assembled via genome alignment. An alternative to both RAST and the CLC Genomics procedure detailed in sections 2.2.7.4 and 2.2.7.5 (Materials and Methods) could be A5-miseq workflow. The A5-miseq pipeline was described as useful for researchers with little bioinformatics experience and produced publishable assemblies, though it is only available on Mac or Linux operating systems. It is an open source, fully automated system which can read long sequences from Illumina MiSeq and assembly genomes up to scaffold level without parameter tuning (Coil, Jospin, & Darling, 2014).

KAAS was used to draw pathways that were not available from Seed Viewer. However it was unable to find and draw xylan/xylose degradation pathway for *L. rhizovicinus* AEOb and *Bacillus licheniformis* AHY1. The starch metabolism diagram included a brief section on xylose utilisation. However, enzymes were present when Brite hierarchies were studied using KAAS.

The majority of bacterial isolates chosen for genome sequencing were thermotolerant. In order to tolerate high temperatures, bacteria have to regulate their fatty acid composition to maintain a highly permeable barrier and liquid crystalline state of the cell membrane (Map of Life, 2012). *L. rhizovicinus* xylanase alignments (Appendix 8.3.10) exhibited a number of polar/non-polar amino acid substitutions. Proteins could be stabilised via amino acid changes, which lead to salt bridge formations and ionic interactions (Madigan & Oren, 1999; Map of Life, 2012). Thermophilic bacterial proteins have highly polar cores, reduced glycine contents, lower surface-to-volume ratios, and a high number of ionic interactions. Polar cores and low surface-to-volume ratios aid folding and stability via solvent exclusion and compaction (Madigan & Oren, 1999). Subtle contributing factors to protein stability also include larger hydrogen bonding and ion pairing networks, as well as smaller loops, lower surface to volume ratio and decreased labile amino acid content. In oligomeric proteins, additional inter-subunit interactions are also important (Epting *et al.*, 2005).

Accumulation of di-glycerol-phosphates have been suggested to aid thermoprotection of proteins (Lamosa *et al.*, 2000; Map of Life, 2012), as well as salt accumulation and increased GC content (Kampmann & Stock, 2004). However, Table 5.1 showed that only six genomes (*L. rhizovicinus* AEOb, AEZa, AEH1 and AFDa, *Burkholderia sp.* AEW2a

and *B. fungorum* AELa) had GC content over 60%, all of which were room temperature isolates that were active at pH 4. Higher GC content stabilises the double helix structure of DNA (Kampmann & Stock, 2004), so it may have been a factor in acid tolerance. Thermosome chaperonins have been reported to bind to heat denatured proteins; averting aggregation and refolding the protein back to their active forms. DNA supercoiling is also thought to be involved in thermotolerance. DNA gyrase, factor for inversion stimulation (FIS) and histone-like nucleoid structuring protein (H-NS) were shown to negatively supercoil and relax circular DNA (Blot, Mavathur, Geertz, Travers, & Muskhelishvili, 2006). Reverse DNA gyrase was shown to positively supercoil DNA which was more resistant to thermal denaturation in hyperthermophilic organisms only (Guipaud, Marguet, Noll, de la Tour, & Forterre, 1997; Madigan & Oren, 1999; Tse-Dinh, Qi, & Menzel, 1997). However, positively and negatively supercoiled plasmids have both been susceptible to thermodegradation (Kampmann & Stock, 2004). DNA gyrase has only been shown to be present only in mesophilic organisms (Guipaud *et al.*, 1997).

A function based comparison was used to attempt to identify acid/alkali/thermo-tolerant genes. Similar genomes were selected and compared with genes only present in the query genome and not the reference. However, the majority of genes involved in extreme tolerance appeared to be up-regulated rather than present/absent (Soemphol *et al.*, 2011). Reverse DNA gyrase, DNA gyrase (EC 5.99.1.3) and chaperone protein DnaK were also identified as thermotolerance/resistance genes. Reverse DNA gyrase has been shown to be present in hyperthermophilic isolates but absent in mesophilic bacteria (Guipaud *et al.*, 1997). Reverse gyrase is used to stabilise DNA via

positive supercoiling (Lowe *et al.*, 1993). DNA gyrase (EC 5.99.1.3) is present in mesophilic bacteria and aids negative supercoiling of DNA, which leads to heat shock stability with relaxed DNA (Tse-Dinh *et al.*, 1997). All of the sequenced genomes with RAST annotations contained a copy of DNA gyrase bar *Bacillus licheniformis* AJF1. However, none of the isolates had annotations for reverse gyrase. Nevertheless, two genomes (*Burkholderia fungorum* AELa and *Burkholderia sp.* AEW2a) had annotations for DNA gyrase inhibitors which could prevent DNA synthesis and are known antimicrobial targets (Boehm *et al.*, 2000; Gellert, O'Dea, Itoh, & Tomizawa, 1976). If these genes were found to be novel they could be investigated in the future. However, a number of *Burkholderia* species have been identified as pathogenic (Du Toit, 2015) and would not be used as a fermentation isolate. They could however, be used for enzyme mining.

Thermotolerance appears to be due to the up-regulation of genes, which have been observed in *Acetobacter tropicalis* SKU1100 (Soemphol *et al.*, 2011). Many genes referred to by Soemphol and colleagues were found in each of the sequenced genomes. RNA-seq could be used to assess whether genes presented by Soemphol were up-regulated in the bacterial isolates within this thesis. RNA-seq could ascertain whether the genes were functional and aided thermo- and acid tolerance. Plus cloning into mesophilic hosts could increase availability of acid/alkali/thermos-tolerant biocatalysts (Adams & Kelly, 1998).

5.4.2 Protein modelling and analysis

Enzymes could only be modelled on proteins that have models/crystal structures when using Phyre2 and 3DLigandSite (Kelley & Sternberg, 2009; Wass *et al.*, 2010). Thus, PSI-Blast pseudo-multiple sequence alignments could hit a 99% identity match but the

model used was only 35% homologous. Such low homology was probably due to the limited availability of proteins, from few bacterial species within the PDB. However, all models used had 100.0 confidence that the probability of the query sequence and the template were homologous. Phyre2 Intensive was therefore used ensure more accurate homology to get a more precise model (Kelley, Mezulis, Yates, Wass, & Sternberg, 2015). Other programmes were available from Phyre2 (One-to-one threading and BackPhyre), however there were time constraints preventing this analysis.

NCBI CDD (Marchler-Bauer *et al.*, 2014) was used for a final check of enzymes before modelling. As Table 5.8 showed; cellulase 1 from *B. subtilis* AJF2 was identified as a zinc peptidase; *L. rhizovicinus* AEOb enzymes hit acetyl esterase/esterase/lipase (xylanase 1) as well as alpha/beta hydrolase (xylanase 2). Cytoplasmic alpha-amylases from *B. coagulans* ABQ were confirmed as alpha-amylases, and the licheninase from *B. subtilis* AJF2 was also confirmed. *B. licheniformis* AHY1 beta-xylosidases 1 and 2 were also confirmed in their identification, as was cellulase 2 from *B. subtilis* AJF2 which also had a cellulose binding domain (CBM).

All three cellulases from *B. subtilis* AJF2 had high homology to previously published sequences from a range of *B. subtilis* isolates (Appendix 8.3.5-8.3.7). As 28 cellulases from *B. subtilis* had been modelled and submitted to RCSB-PDB (Research Collaboratory for Structural Bioinformatics Protein Data Bank), this was expected. Thus models were thought to be relatively accurate. *B. licheniformis* AHY1 beta-xylosidases were also highly homologous to top PSI-BLAST hits (Appendix 8.3.13). However, beta-xylosidase 1 only had 25% identity with the top model chosen and beta-xylosidase 2

had only 35% identity. Both enzymes were modelled using the same crystal structure from RCSB-PDB (1YIF, beta-1,4-xylosidase from *B. subtilis*) which was classed as homo 4-mer-A4. Four copies of this enzyme would therefore be required to form the entire complex (Patskovsky, Y., Almo, 2009).

Insufficient homologs were found to model ligand binding sites for *L. rhizoviciuns* AEOb xylanase 1 so no predictions were made. However, when modelled in Phyre2 intensive, confidence was high (100%) with 78% coverage (Table 5.9). Alignment of xylanase 2 found in *L. rhizovicinus* AEOb had less homology to the PSI-BLAST top hit B3VP83 (Appendix 7.3.11) with 20.1% identity and 1e-55. When enzymes from *L. rhizovicinus* AEOb were aligned using NCBI BLAST homology was relatively high. Xylanse 1 had query coverage of 94%, e-value of 4e-141, and maximum identity of 72% to hypothetical protein from *Luteibacter sp.* 9133. As hypothetical proteins were the top hits from NCBI BLAST it was unsurprising that ligand binding sites could not be modelled. Xylanase 2 had query coverage of 96%, e-value of 1e-106, and identity of 60% to xylanase from *Caulobacter sp.* JGI 0001010-J14. However, when *L. rhizovicinus* DSM 16549 was analysed, high homology was found in equivalent xylanalytic enzymes (Appendix 8.3.10-8.3.12). Nevertheless, these annotations were not present on the GenBank genome.

Both cytoplasmic alpha-amylases from *B. coagulans* ABQ had high homology with top hits from PSI-BLAST (amylase 1: 62.2% identity, e-value 3e-73; amylase 2: 46.7% identity, e-value 1e-75). However amylase 2 had only 33% identity with the top template (3AIC, (Ito *et al.*, 2011)) used for modelling in Phyre2 Intensive, whereas amylase 1 had 51% identity with the top model (1HVX, (Suvd, Fujimoto, Takase,

Matsumura, & Mizuno, 2001)). However neither query sequence was modelled on *B. coagulans* proteins due to only three proteins from this strain were deposited in RCSB-PBD (RCSB-PDB, 2005). Though amylase 1 was modelled on a thermostable alpha-amylase (Suvd *et al.*, 2001).

Binding site predictions from 3DLigandSite were less successful than anticipated. Many modelled enzymes had very few residues predicted for ligand binding (Table 5.9). Many of the binding sites shown were metallic heterogen binding sites. Binding metallic heterogens has been shown to aid efficiency of the enzyme. Calcium chloride has been shown to be a critical factor for alpha-amylase production and activity (Chakraborty *et al.*, 2011) and to increase thermostability of alpha-amylase (Chimata *et al.*, 2011). However, some alpha-amylase proteins did not require Ca^{2+} to be acidophilic and thermostable (Sharma & Satyanarayana, 2011). Also, enhanced activity by magnesium (Mg^{2+}) and calcium (Ca^{2+}) with xylanases has been reported. Zinc had been reported to both activate and inhibit xylanase proteins (Khasin *et al.*, 1993; Pal & Khanum, 2010; van Dyk *et al.*, 2010; Y. Yang *et al.*, 2010).

5.4.3 Conclusions

Whole genome sequencing was used to confirm or disprove the 16S rRNA based identity of bacterial isolates and determine any novel isolates or subspecies. It also aided enzyme identification via bioinformatics analysis, and identified where gene transfer had potential to occur. Whole genome sequencing and protein modelling was used to potentially identify two putative novel xylanases. Also, a potential subspecies (*L. rhizovicinus* AEOb, AEZa and AFDa) was discovered. A novel genome, for which only the 16S rRNA sequence could be relied upon for identification, was also identified as *S. assacharolytica* AEH1. No current genome was available for comparison for AEH1.

Pathway analysis was executed with RAST and KAAS which could be used to design metabolic engineering approaches. Whereby, whole pathways could be engineered into model or other species. Further work was required to characterise extremophile changes in enzymes (discussed in section 6.4). Protein modelling could aid synthetic biology and aid further understanding of each enzyme, plus potentially view any acid/alkali/thermo-tolerance changes through homology (Bergquist *et al.*, 1999; Madigan & Oren, 1999).

Chapter 6 Discussion, Conclusions and Future Work

6.1 Aims and Background

The main aim of this project was to isolate a range of bacteria, with genes encoding enzymes useful for saccharification of plant cell wall and starch polymers, from pre-treated biomass. Isolates from *Miscanthus* chip and plots were characterised using carbon utilisation assays based on such polymers (Chapter 3). Results from characterisation aided selection of isolates for whole genome sequencing in order to understand depolymerising enzymatic action (Chapter 5). Characterisation of the alpha-amylase protein (Chapter 4) aided understanding of enzyme kinetics and established strand of study in which enzymes could be taken from protein modelling. The research detailed within this thesis lay down a foundation for future projects and discoveries. Not only within the biorefining area, but also into other aspects of microbiology, such as antimicrobials, baking, brewing, paper bleaching, and other industries. Without this ground work and discoveries, future endevours may not be possible.

6.2 Key Findings and Conclusions

- Isolation and Characterisation of the Bacterial Library
    - 148 bacteria were isolated and identified using multiple primer pairs for 16S rRNA gene sequencing, as not all universal primers were universal. Of these isolates, 45 were identified to a genus level and 103 to a species level. The most commonly occurring isolates were *Burkholderia fungorum*, *Psuedomonas spp.* and *Bacillus spp.*. Ten isolates identity were changed using whole genome sequencing, due to closely related species that were not differentiated using 16S rRNA sequencing.

- A wide range of xylanase, cellulase and starch degrading activities at a range of temperatures (room temperature to 70 °C) and pH values (4-11) were observed. Specifically, 88.5% of cultures showed cellulose activity, 93.2% xylanase activity, and 79.7% starch activity. Further to this, 66.2% of isolates had activity across all three carbon sources. Bacterial isolates were screened for activity at extremes of pH analogous to steam explosion and AFEX pre-treatment. Activity at the extremes showed that 5.4% of isolates had xylan activity at pH 4 and 15.5% at pH 11, 4.7% of these activitys were present at 60 °C (*B. subtilis* AHU, AIL (did not amplify), AJF2, AJK; and *B. licheniformis* AHY2, AJG, and *Unknown sp.* AIS1b, which had erratic activity). No cellulase activity was observed at pH 4, though 10.1% of isolates had activity at pH 10, with only 2% of these (*B. licheniformis* AIB (did not amplify), *B. subtilis* AJN1a, and *Unknown sp.* AIU, which had erratic activity) at 60 °C. Activity on starch at pH 4 was observed for 18.2% of isolates, and 23.6% at pH 9, though only 1.4% occurred at this pH at 60 °C (*B. licheniformis* AJF1 and *B. subtilis* AJN1a). The bacterial isolates that had consistent activity at high pH and temperature were chosen for whole genome sequencing. Bacterial growth and enzymatic activity were observed at 70 °C, though optimal temperatures were 40-50 °C for thermotolerant species

- Enzyme Isolation and Expression
    - Novel discovery of *B. coagulans* ABQ alpha-amylase double pH optima, which has not been observed in *B. coagulans* previously. However,

fungal (Igbokwe, Ngobidi, & Iwuchukwu, 2013), mollusc (Hameed & Paulpandian, 1987), and *Bacillus licheniformis* (Ramesh & Lonsane, 1989) alpha-amylases have been reported to have two optimal pHs. Double pH optima have previously been observed in polyphenol oxidases, as well as modified xylanase *LAX* from *Scopulariopsis sp.* (Afzal, Bokhari, & Siddiqui, 2007).

    - Confirmation of glucanase secretion into the media.

- Next Generation Sequencing of Miscanthus Associated Bacteria and Protein Modelling of Sugar Releasing Enzymes
    - 23 genomes were sequenced and assembled into contigs, annotated and searched for homology to known enzymes. From this, key enzymes involved in plant cell wall saccharification were identified (xylanase, beta-xylosidase, alpha-amylase, cellulose and licheninase) via pathway analysis using RAST and KAAS.
    - Potentially novel bacterial subspecies within the sequenced genomes were described; one of which had three independent clones (*Luteibacter rhizovicinus* AEOb, AEZa, and AFDa, and *Sphingomonas asaccharolytica* AEH1, Figure 5.4). Differences in genome homology to reference strains, could be due to adaptation to the environment they were isolated from.
    - Two unpublished xylanases (within the published reference but not annotated) were identified by homology in *L. rhizovicinus* AEOb. Xylanase 2 appeared to be a novel due to negligible homology at the C

terminus. However, sequencing error could have caused such differences.

6.3 General Discussion

Considering the original BEACON pipeline, the research detailed within this thesis comes under secondary processing. However, due to the findings described pertaining to sugar release, this work could bridge the gap from secondary processing to conversion. Synthetic biology could also help to speed this process up, by combining subunits from enzymes that have extremophilic tendencies, to create an engineered enzyme that has a high depolymerisation rate and can withstand pre-treatment conditions. Such subunit optimisation and enzyme modification could also aid efficiency, by removing inhibitor succeptability.

At present there are no publications detailing isolation of bacterial species from bioenergy grass *Miscanthus* bar nitrogen-fixing endophytes (Eckert *et al.*, 2001; Kirchhof *et al.*, 2001). Therefore, this is the first study to isolate, identify and characterise aerobic bacteria from degrading *Miscanthus* material. Generation of a large library of bacteria, isolated from *Miscanthus* chip and soil surrounding *Miscanthus* plots, increased the probability that novel species or enzymes would be found. It also increased the chance that species useful for saccharification and biorefinery of pre-treated biomass would be isolated, as bacteria would have adapted to their environment. Isolates could therefore have genes encoding enzymes specific for sugar release from *Miscanthus* (Kongpol *et al.*, 2009; Morales *et al.*, 1993; Reinhold-Hurek & Hurek, 2011; Tse-Dinh *et al.*, 1997; van Dijl & Hecker, 2013). Species screened at mesophilic temperatures across a pH range provided enzyme activities, which could pre-empt yeast fermentation without raising costs by the need for cooling

(Abdel-Banat *et al.*, 2010). Screening thermotolerant isolates also meant that enzyme activities could be observed with a view for alternative fermentations, such as anaerobic digestion, *Clostridium* fermentation or solid state fermentation. Such fermentation systems were reported to be suited for higher temperatures (Angelidaki & Ahring, 1993; Asha Poorna & Prema, 2007; Bayané & Guiot, 2010; Chen, Cheng, & Creamer, 2008; Freier *et al.*, 1988; Mata-Alvarez, Macé, & Llabrés, 2000; Nakayama *et al.*, 2011; Qureshi *et al.*, 2014; The Wales Centre of Excellence for Anaerobic Digestion, 2008). The pH screen at high temperatures enabled enzymatic activities to be assessed, whereby within fermentation, acidic metabolic products would be produced thus lowering pH, or if alkaline pre-treatments were required. Therefore, the enzymatic activity range could be gauged for each isolate from acidic to alkaline conditions. Though bacteria were isolated from mesophilic conditions (soil and *Miscanthus* chip in the field), a number of isolates were thermotolerant, and had enzymatic activity across the carbon utilisation assays. The majority of thermotolerant bacteria were *Bacillus sp.* and have therefore been reported to have a wide range of physiological diversity for temperature and pH (Holt *et al.*, 2000). However, temperature has been shown to increase enzymatic activity (until denaturation) due to faster molecules colliding more frequently (Stoker, 2012), which could explain activity at 70 °C though little bacterial growth was observed.

The use of Gram's iodine to stain plates for carbon utilisation profiles was chosen for its quick and clear results (Maki *et al.*, 2009). The pH range for carbon utilisation profiles was devised by assessment of what was known about each class of enzyme (

Table 2.6). Both bacterial growth and activity and pure enzymatic activity were taken into consideration for the pH range, as some enzymes were secreted into the media. Secreted enzymes, as demonstrated in section 4.2.8, would therefore be exposed to the acidic or alkaline environment.

Bacillus coagulans ABQ was investigated thoroughly, as *B. coagulans* have been reported as a lactic acid producers and could therefore have genes encoding useful enzymes for saccharification (Payot *et al.*, 1999). The alpha-amylase isolated from this strain was pursued due to its extracellular nature, as it was both soluble and insoluble and contained a signal peptide. The protein could be of use in single-step fermentations, such as those exhibited by *Lactobacillus sp.* (Reddy, Altaf, Naveena, Venkateshwar, & Kumar, 2008). Sequencing the whole genome of ABQ allowed the starch metabolism pathway to be identified (Figure 5.9) and this could be explored in greater detail in the future. The genomes of 22 other isolates were sequenced, due to their enzymatic activity at low pH at mesophilic temperature and high pH at thermophilic temperatures (section 3.2). In this way, potentially novel pathways and enzymes could be identified with a view to metabolically engineer them into model species. Of the 23 genomes sequenced, four could not be identified by the closest neighbour reported by Seed-RAST and alignment of the largest contig via NCBI BLAST. However, they were identified by 16S rRNA gene found within the genome rather than PCR amplification. Of these four a single *L. rhizovicinus* (AEOb) was used for protein discovery. The enzymes analysed from AEOb had less than 97% query coverage and low homology to other proteins in NCBI BLASTp database (Table 5.5 and 5.6).

However, alignment with the published genome highlighted highly homologous regions (Appendix 8.3.14).

Eight main traits were observed within the five prominent clades of the 16S rRNA phylogenetic tree (Figure 3.4) for the bacterial isolate library, based on published data of homologous species. Homologous sequences were identified as: plant pathogenicity; biotechnologically useful; of use for enzyme production in industry; of use in biorefining; promotion of plant growth; endophytic relationship with plants; symbiotic relationship with plants, and human pathogenicity. A number of these traits were expected, such as plant pathogenicity, growth promotion, symbiosis and endophytic relationships, as isolates were cultured from degrading *Miscanthus* chip or plots. Human pathogenic strains were included due to the number of isolates that could only be identified to a genera level, and could potentially include *Burkholderia cepacia*, *Chryseobacterium indologenes*, *Pseudomonas aeruginosa*, *Serratia grimesii*, and *Staphylococcus aureus* (Berg *et al.*, 2005). However, such pathogens cover a range of traits, including plant pathogenicity, and could therefore be a useful source of enzymes for biorefinery.

The bacterial library spanned 42 genera. However, the majority of variation came from mesophilic isolation whereas only four genera were present in the thermophiles, with six isolates unknown. The majority of activity at higher alkalinity was observed within the thermophiles. However, the only activity seen at pH 4 was observed within the mesophiles. This could be due to the fact that pH rose with temperature (Table 2.45) so pH measured at 10 when making the agar for carbon utilisation profiles would

actually be pH 10.45 at 50 °C and maybe higher at 60-70 °C (Barron, Ashton, & Geary, 2006).

Overall characterisations of the bacterial library led to selection for whole genome sequencing. However, *B. coagulans* ABQ was also included in selection for whole genome sequencing due to alpha-amylase characterisation and the discovery of two pH optimas (Chapter 4). Sequencing the whole genome allowed the starch degradation pathway to be isolated and studied for ABQ, and other pathways (xylose utilisation and cellulose degradation) for isolates *Unknown sp.* AEOb, *B. licheniformis* AHY1 and *B. subtilis* AJF2. The initial enzyme required for starch depolymerisation (to dextrin) is alpha-amylase. An alpha-amylase from *B. coagulans* ABQ was amplified, cloned and expressed in *E. coli* (Chapter 4). A large quantity of functional protein was produced and characterised, though assays were not optimised due to time constrains. However, assays showed that there were two optimal pH activity peaks for the alpha-amylase, which has not been observed previously for this class of enzyme, but is known to be present in polyphenol oxidases (Saeidian & Rashidzadeh, 2013). However, ABQ had two cytoplasmic alpha-amylases, with different protein sequences and different active sites (section 8.3.13). Alpha-amylase 2 however, had 92% identity to other *B. coagulans* sequences within the NCBI database when the amino acid sequence was aligned, whereas alpha-amylase 1 had 99% homology. Future work could include optimisation and further characterisation of amylase 1 as well as characterisation of amylase 2 as this protein may have greater activity.

Enzymes were modelled, based on protein sequences identified by homology, from whole genome sequenced isolates. Modelled enzymes were selected based on the

isolates performance, at high or low pH in the carbon utilisation assays (Chapter 3), and the proposed functionality of the pathway (i.e. a full pathway was present). Though proteins were modelled on homology to existing models, which were based on crystal structures, binding sites could not be modelled accurately. This was probably due to the fact that models were based on entries in PDB, which contains crystal structers from limited species. Two xylanases were modelled from potentially new subspecies *L. rhizovicinus* AEOb, though Xylanase 1 could not be modelled in 3DLigandSite, due to insufficient homologous structures. Xylanase 1 could therefore be presumed to have a novel structure, though the sequence was homologous to a potential xylanase in the DSM 16549 genome. Xylanase 2 could be modelled, but had large regions of low homology to its counterpart in the DSM 16549 genome. AEOb had enzymatic activity on xylan agar from pH 4 to 8.5. Potentially xylanase 1 and 2 could have adapted to acidic conditions, and therefore folded in a different confirmation (Goto, Calciano, & Fink, 1990; Jaenicke & Závodszky, 1990; Jaenicke, 1991; Sasahara, Demura, & Nitta, 2002; Szila, Kardos, Osvath, Barna, & Zavodszky, 2007; Tehei, 2005). Both xylanases were unpublished, with only RAST homology identifying them, and xylanase 2 is a novel find.

Targeting bacteria and enzymes for higher efficiency and tolerance of environments, required for pre-treatment of biomass, compared to a synthetic approach could be useful for characteriation. However, when dealing with bacteria that could be potential human pathogens, a synthetic approach would be more advisable. However, this project could progress biorefinery methods by eliminating wash steps from pre-treated biomass. Enzymes or bacterial isolates could be used for saccharification while

biomass is still at a high or low pH, especially for maximum sugar realease from both cellulose and hemicellulose. Alternatively, isolates that were found to have activity at high temperatures could be utilised for sugar release, without controlled conditions. Novel enzymes presented in this project could be of use in biorefining to maximise sugar release, and could cut costs of enzyme production, as well as temperature and pH control.

6.4 Future Work

To isolate a greater range of bacteria other types of media could be utilised, such as Tryptone Soya agar. Specific carbon sources could be integrated into media for a more targeted approach. Since only aerobic isolation was performed a plethora of species were missed. Anaerobic isolation could yield a number of useful species. Isolation from a greater range of samples such as soils, *Miscanthus* chip, high-sugar rye grass, and composted materials could improve range of species. Once bacterial species were isolated, a range of temperatures could be used to predict optimal growth conditions. Carbon utilisation profiling, growth curve analysis of interesting isolates, and HPLC analysis of supernatant over time, would provide a more clear understanding of sugar release and product formation per isolate, in a biorefining context. In addition, a modification could be made to the Lee *et al.* (2012) high-throughput biological conversion assay, whereby *Miscanthus* and other biomass crop cell-wall extractions were used as a carbon source. Bacterial isolates could be inoculated into cell-wall extractions, for conversion rates of lignocellulosic material to ethanol (S. J. Lee, Warnick, Leschine, *et al.*, 2012). Using plant cell wall extractions would therefore give a more realistic picture of how each isolate would behave in a biorefining context. Finally, fosmid libraries could be generated, as well as wider isolation techniques.

Fosmid libraries could aid the isolation of novel enzymes, which would not be picked up in whole genome sequencing, if there was little homology with known proteins. Such libraries could also be used to assay for antimicrobials.

The carbon utilisation profiles of mesophiles should be screened across a temperature range also. A high throughput, two stage system should be implemented, whereby isolates would be screened on agar using a hedgehog (96 Solid Pin Multi-Blot Replicator) across a pH range, for a shorter period of time (e.g. 8 hours). This would give a presence/absence halo result, which would be followed by a full screen of those with activity, across a full pH range (e.g. pH 4, 5, 6, 7, 8, 9, 10, 11, 12) and temperature range (e.g. 30-80 °C in 5 °C increments). Such techniques could also be used, if fosmid/clone libraries were generated for enzymatic activity screening (Y. S. Kim *et al.*, 2000).

Isolation, cloning and expression of enzyme coding genes followed by characterisation of the expressed protein were very useful. Such techniques could be employed for a range of enzymes, especially those characterised in Chapter 5 (section 5.3.6). By expressing these proteins, functionality could be confirmed and kinetics established. Also, further characterisation of the alpha-amylase enzyme (such as V_{max} and K_m), and optimisation of each experiment could result in a more active profile, and create publishable data on its novel double pH optima. Future work could include excision of protein bands, from SDS-PAGE, followed by mass spectrometry to identify the amino acid sequence. Protein sequences would then be used to find homologous genes, via NCBI database searches (Sánchez-Herrera *et al.*, 2007; Westereng *et al.*, 2011), as well as confirmation of potential genes in the sequenced genomes. Enzyme discovery and

following characterisation would therefore render novel enzymes, whereas the main approach described in Chapter 4 was using both nucleotide and amino-acid sequence based homology. Also, the use of mass-spectrometry for metabolic products would aid identification of unknown enzymes (Gool, 2012). Isolation of bacteria, with selection pressures such as acid/alkali/thermo- tolerances, could then feed directly into isolation of novel enzymes due to environment adaption. Therefore, rather than isolating a range of bacteria then subjecting them to different environments to test enzymes; instead enzymes would already be selected for, and greater characterisation of these potentially novel proteins could occur.

A more efficient way to find genes related to acid/alkali/thermo-tolerance would be to write a script. Such a script could then filter against known genes, to find homology for novel carbohydrate degrading and any extremophile genes. Using a script would cut time constraints. It would much quicker and more efficient than manually searching or using BLAST to find genes against one genome at a time.

A more intense approach to protein modelling, from an amino acid sequence, could involve Phyre2 One-to-one threading and Phyre2 BackPhyre (using Phyre2 Expert Mode). One-to-one threading aligns a query sequence to a particular known structure, via threading the sequence through the model. A model is then constructed with more accurate folds and turns. BackPhyre could be used to find homologous structures in a range of genomes, from a user input of a modelled protein (Kelley *et al.*, 2015). Mauve 2.4.0 software (Darling, Mau, Blattner, & Perna, 2004) could be used to align a number of genomes for comparison, to identify homology and missing genes. The software

could also be used to assemble genomes based on a reference. All three programmes could be utilised to maximise data on the 23 sequenced genomes.

Synthetic biology and metabolic engineering, for high yields of biofuels or products, could utilise genome sequencing and protein modelling. Next generation sequencing of genomes could identify fully active enzymatic pathways, which could then be metabolically engineered it into model species, such as *Saccharomyces cerevisiae* or *Escherichia coli.* Pathways could be more easily targeted when the sequence is known, and if the proteins are located near to each other, and under the control of the same promoters (Connor & Atsumi, 2010; Medema *et al.*, 2011; F. Zhang *et al.*, 2011). Whole genome sequencing techniques could also be used to identify genetic characteristics, and improve the uses of bacterial species such as *B. coagulans* (Su & Xu, 2014). By aligning a number of whole genome sequences for strains, difference could be observed, especially due to adaptation to e.g. acidic or alkaline conditions (Tehei, 2005). Alternatively, a pathway of genes could be compared to already published data; differences could indicate improved or loss of function. After characterising the isolates activity on the required substrate, the pathway could be metabolically engineered into another strain or species. A cell free production of the protein cascade could be economically utilised, so that only specific products are yielded, without unwanted metabolic byproducts (Carlson, Gan, Hodgman, & Jewett, 2012).

Protein modelling has been shown to work and could potentially advance biofuel production. Xylanase (*XynA*) has been redesigned (based on endo-xylanase via YASARA software) into chimeric form, and shown to be more active (Santos *et al.*, 2014).

Therefore, if a range of enzymes were available within PDB, on which to model proteins accurately via crystal structures, redesigned enzymes could be produced. Synthetic enzyme mixtures (accessory and core enzymes used to degraded biomass pre-treated by AFEX) (Banerjee, Car, *et al.*, 2010a, 2010b), and cell free enzyme cascades, have been shown to be economically viable for biofuel production, due to high production of low value biofuels and low complexity in the biological system (Y. H. P. Zhang, Sun, & Zhong, 2010). Synthetic enzyme supercomplexes could be used to create metabolic pathways that do not occur in nature, from a wide range of organisms with different characteristics. Each stage of degradation could be controlled by immobilisation of enzymes, which also aids enzyme recovery, avoids unfavourable kinetics, and enhances performance via substrate channelling (Ferrer *et al.*, 2007; Kazenwadel, Franzreb, & Rapp, 2015). Therefore, proteins from the xylan degradation pathway, from *L. rhizovicinus* AEOb or *B. licheniformis* AHY1, could be isolated and combined to break down hemicellulose, at range of temperatures and pH. Alternatively, engineered *E. coli* isolates could be used to express each enzyme in a cascade (Shin *et al.*, 2010). There is also potential to design whole nanomachines (cellulosomes), to degraded lignocellulosic material for biofuel production (Fontes & Gilbert, 2010).

From the genomes reviewed in Chapter 5, 10 sequenced isolates were annotated using RAST to have xyloside transporter XynT (*B. pumilus* AIF, *B. licheniformis* AJG, *B. subtilis* AHU, *B. licheniformis* AIB, *B. subtilis* AJF2, *Unknown sp.* AEH1, *B. coagulans* ABQ, *B. subtilis* AJK, *B. pumilus* AIK, *and B. subtilis* AJN1a). Xyloside transporters (e.g. XynT) have been transformed into yeast or *E. coli* and been active (Fernandes &

Murray, 2010; Shin *et al.*, 2010). Thus, genome sequencing could highlight a range of transporters that have homology, but would not necessarily be identified via PCR.

Finally, transcriptome expression analysis such as RNAseq would, give a clearer picture of which enzymes were expressed, in what order, in response to which environment. Novel enzymes could also be discovered in this manner (Qi *et al.*, 2011). Also, analysis of genomes both unidirectionally and bidirectionally would avoid bias and give a more complete picture of each genome (Cock & Whitworth, 2010).

Chapter 7 References

Abdel-Banat, B. M. a, Hoshida, H., Ano, A., Nonklang, S., & Akada, R. (2010). High-temperature fermentation: How can processes for ethanol production at high temperatures become superior to the traditional process using mesophilic yeast? *Applied Microbiology and Biotechnology*, *85*(4), 861–867. doi:10.1007/s00253-009-2248-5

Abraham, E. G., Nagaraju, J., & Datta, R. K. (1992). Biochemical studies of amylases in the silkworm, Bombyx mori L.: Comparative analysis in diapausing and nondiapausing strains. *Insect Biochemistry and Molecular Biology*, *22*(8), 867–873. doi:10.1016/0965-1748(92)90113-S

Adams, M. W. W., & Kelly, R. M. (1998). Finding and using hyperthermophilic enzymes. *Trends in Biotechnology*, *16*(8), 329–332. doi:10.1016/S0167-7799(98)01193-7

Adapa, P. K., Engineering, B., Karunakaran, C., Tabil, L. G., & Schoenau, G. J. (2009). Qualitative and Quantitative Analysis of Lignocellulosic Biomass using Infrared Spectroscopy. In *CSBE/SCGAB 2009 Annual Conference* (pp. 1–21). Rodd's Brudenell River Resort, Prince Edward Island: The Canadian Society for Bioengineering.

Adger, W. N., Huq, S., Brown, K., Conway, D., & Hulme, M. (2003). Adaptation to climate change in the developing world. *Progress in Development Studies*, *3*(3), 179–195. doi:10.1191/1464993403ps060oa

Advisory Committee on Dangerous Pathogens. (2013). *The Approved List of Biological Agents*. *Health & Safety Executive*. London.

Afzal, A. J., Bokhari, S. A., & Siddiqui, K. S. (2007). Kinetic and Thermodynamic Study of a Chemically Modified Highly Active Xylanase from Scopulariopsis sp . A HMED J. *Applied Biochemistry And Biotechnology*.

AgarGel. (2003). Agar-Agar Properties and Specifications. Retrieved from http://www.agargel.com.br/agar-tec-en.html

Agilent Technologies. (2005). EZChrom Elite Installation Guide. Santa Clara: Scientific Software Inc.

Agler, M. T., Wrenn, B. a, Zinder, S. H., & Angenent, L. T. (2011). Waste to bioproduct conversion with undefined mixed cultures: the carboxylate platform. *Trends in Biotechnology*, *29*(2), 70–8. doi:10.1016/j.tibtech.2010.11.006

Aguilar, R., Ramirez, J. A., Garrote, G., & Vazquez, M. (2002). Kinetic study of the acid hydrolysis of sugar cane bagasse. *Journal of Food Engineering*, *55*, 309–318.

Ahmad, M., Roberts, J. N., Hardiman, E. M., Singh, R., Eltis, L. D., & Bugg, T. D. H. (2011). Identification of DypB from Rhodococcus jostii RHA1 as a lignin peroxidase. *Biochemistry*, *50*(23), 5096–107. doi:10.1021/bi101892z

Akinterinwa, O., Khankal, R., & Cirino, P. C. (2008). Metabolic engineering for bioproduction of sugar alcohols. *Current Opinion in Biotechnology*, *19*(5), 461–7. doi:10.1016/j.copbio.2008.08.002

Albers, S. V., Koning, S. M., Konings, W. N., & Driessen, a. J. M. (2004). Insights into

ABC Transport in Archaea. *Journal of Bioenergetics and Biomembranes*, *36*(1), 5–15. doi:10.1023/B:JOBB.0000019593.84933.e6

Allen, C., & Pinon, D. (2013). Syngenta and Bonanza BioEnergy Reach Commercial Agreement to Use Enogen Trait Technology in 2013. Syngenta. Retrieved from www.syngenta-us.com/news_releases/NewsPrint.aspx?id=170456

Allgaier, M., Reddy, A., Park, J. I., Ivanova, N., D'haeseleer, P., Lowry, S., … Hugenholtz, P. (2010). Targeted discovery of glycoside hydrolases from a switchgrass-adapted compost community. *PloS One*, *5*(1), e8812. doi:10.1371/journal.pone.0008812

Allison, G. G. (2011). Application of fourier transform mid-infrared spectroscopy (FTIR) for research into biomass feed-stocks. In G. Nikolic (Ed.), *Fourier Transforms – New Analytical Approaches and FTIR Strategies* (First., pp. 71–88). Rijeka: InTech. doi:10.5772/2040

Allison, G. G., Morris, C., Clifton-Brown, J., Lister, S. J., & Donnison, I. S. (2011). Genotypic variation in cell wall composition in a diverse set of 244 accessions of Miscanthus. *Biomass and Bioenergy*, *35*(11), 4740–4747. doi:10.1016/j.biombioe.2011.10.008

Almarsdóttir, A. R. (2011). *Thermophilic ethanol and hydrogen production from lignocellulosic biomass*. University of Akureyri.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. doi:10.1016/S0022-2836(05)80360-2

Alves, L. a., Felipe, M. G. a., Silva, J. B. A. E., Silva, S. S., & Prata, A. M. R. (1998). Pretreatment of sugarcane bagasse hemicellulose hydrolysate for xylitol production byCandida guilliermondii. *Applied Biochemistry and Biotechnology*, *70-72*(1), 89–98. doi:10.1007/BF02920126

Amábile-Cuevas, C. F., & Chicurel, M. E. (1993). Horizontal gene transfer. *American Scientist*, *81*(4), 332–341.

Andreolli, M., Lampis, S., Zenaro, E., Salkinoja-Salonen, M., & Vallini, G. (2011). Burkholderia fungorum DBT1: a promising bacterial strain for bioremediation of PAHs-contaminated soils. *FEMS Microbiology Letters*, *319*(1), 11–8. doi:10.1111/j.1574-6968.2011.02259.x

Angelidaki, I., & Ahring, B. K. (1993). Thermophilic anaerobic digestion of livestock waste: the effect of ammonia. *Applied Microbiology and Biotechnology*, *38*(4), 560–564. doi:10.1007/BF00242955

Angelov, A., Loderer, C., Pompei, S., & Lieb, W. (2011). Novel family of carbohydrate-binding modules revealed by the genome sequence of spirochaeta thermophila DSM 6192. *Applied and Environmental Microbiology*, *77*(15), 5483–5489. doi:10.1128/AEM.00523-11

Angenent, L. T., & Wrenn, B. a. (2008). Optimising mixed-culture bioprocessing to convert wastes into bioenergy. In J. D. Wall, C. S. Harwood, & A. Demain (Eds.), *Bioenergy* (pp. 179–194). Washington DC: ASM Press.

Anthony, C. (1975). The Microbial Metabolism of C1 Compounds. *Biochemical Journal*, *146*, 289–298.

Artursson, V., & Jansson, J. K. (2003). Use of Bromodeoxyuridine Immunocapture to Identify Active Bacteria Associated with Arbuscular Mycorrhizal Hyphae. *Applied and Environmental Microbiology*, *69*(10), 6208–6215. doi:10.1128/AEM.69.10.6208-6215.2003

Asgher, M., Asad, M. J., Rahman, S. U., & Legge, R. L. (2007). A thermostable α-amylase from a moderately thermophilic Bacillus subtilis strain for starch processing. *Journal of Food Engineering*, *79*(3), 950–955. doi:10.1016/j.jfoodeng.2005.12.053

Asha Poorna, C., & Prema, P. (2007). Production of cellulase-free endoxylanase from novel alkalophilic thermotolerent Bacillus pumilus by solid-state fermentation and its application in wastepaper recycling. *Bioresource Technology*, *98*(3), 485–90. doi:10.1016/j.biortech.2006.02.033

Ashelford, K. E., Chuzhanova, N. A., Fry, J. C., Jones, A. J., & Weightman, A. J. (2005). At Least 1 in 20 16S rRNA Sequence Records Currently Held in Public Repositories Is Estimated To Contain Substantial Anomalies. *Applied and Environmental Microbiology*, *71*(12), 7724–7736. doi:10.1128/AEM.71.12.7724

ATCC. (2013). *(ATCC® BAA 1015™) Luteibacter rhizovicinus product sheet*. Manassas. Retrieved from www.atcc.org/~/ps/BAA-1015.ashx

Attotron Biosensor Corporation. (2014). Transcrition and translation Tool. Retrieved March 10, 2015, from http://www.attotron.com/cybertory/analysis/trans.htm

Aziz, R. K., Bartels, D., Best, A. a, DeJongh, M., Disz, T., Edwards, R. a, … Zagnitko, O. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, *9*, 75. doi:10.1186/1471-2164-9-75

Bahadure, R. B., Agnihotri, U. S., & Akarte, S. R. (2010). Assay of population density of amylase producing bacteria from different soil samples contaminated with flowing effluents, *2*(1), 9–13.

Bakker, E. P. (1990). The role of alkali-cation transport in energy coupling of neutrophilic and acidophilic bacteria: An assessment of methods and concepts. *FEMS Microbiology Letters*, *75*(2-3), 319–334. doi:10.1111/j.1574-6968.1990.tb04105.x

Ballschmiter, M., Armbrecht, M., Ivanova, K., Antranikian, G., & Liebl, W. (2005). AmyA, an a-Amylase with b-Cyclodextrin-Forming Activity, and AmyB from the Thermoalkaliphilic Organism Anaerobranca gottschalkii: Two α-Amylases Adapted to Their Different Cellular Localizations. *Applied and Enviromental Microbiology*, *71*(7), 3709–3715. doi:10.1128/AEM.71.7.3709

Banerjee, G., Car, S., Scott-Craig, J. S., Borrusch, M. S., Aslam, N., & Walton, J. D. (2010a). Synthetic enzyme mixtures for biomass deconstruction: production and optimization of a core set. *Biotechnology and Bioengineering*, *106*(5), 707–20. doi:10.1002/bit.22741

Banerjee, G., Car, S., Scott-Craig, J. S., Borrusch, M. S., Bongers, M., & Walton, J. D. (2010b). Synthetic multi-component enzyme mixtures for deconstruction of lignocellulosic biomass. *Bioresource Technology*, *101*(23), 9097–105. doi:10.1016/j.biortech.2010.07.028

Banerjee, G., Car, S., Scott-Craig, J. S., Borrusch, M. S., & Walton, J. D. (2010c). Rapid optimization of enzyme mixtures for deconstruction of diverse pretreatment/biomass feedstock combinations. *Biotechnology for Biofuels*, *3*(1), 22. doi:10.1186/1754-6834-3-22

Banerjee, G., Car, S., Scott-Craig, J. S., Hodge, D. B., & Walton, J. D. (2011). Alkaline peroxide pretreatment of corn stover: effects of biomass, peroxide, and enzyme loading and composition on yields of glucose and xylose. *Biotechnology for Biofuels*, *4*(1), 16. doi:10.1186/1754-6834-4-16

Banerjee, G., Scott-Craig, J. S., & Walton, J. D. (2010). Improving Enzymes for Biomass Conversion: A Basic Research Perspective. *BioEnergy Research*, *3*(1), 82–92. doi:10.1007/s12155-009-9067-5

Bano, S., Ul Qader, S. A., Aman, A., Syed, M. N., & Azhar, A. (2011). Purification and characterization of novel α-amylase from Bacillus subtilis KIBGE HAS. *AAPS PharmSciTech*, *12*(1), 255–61. doi:10.1208/s12249-011-9586-1

Bansod, S. M., Dutta-Choudhary, M., Srinivasan, M. C., & Rele, M. V. (1993). Xylanase active at high pH from an alkalotolerant Cephalosporium species. *Biotechnology Letters*, *15*(9), 965–970. doi:10.1007/BF00131765

Barley, S. (2010). Where there's bugs, there's brass British firm lands $500m biofuel contract Environment The Guardian. Retrieved from http://www.guardian.co.uk/environment/2010/sep/20/tmo-renewables-bacteria-biofuel-fiberight

Barnette, A. L., Bradley, L. C., Veres, B. D., Schreiner, E. P., Park, Y. B., Park, J., … Kim, S. H. (2011). Selective detection of crystalline cellulose in plant cell walls with sum-frequency-generation (SFG) vibration spectroscopy. *Biomacromolecules*, *12*(7), 2434–9. doi:10.1021/bm200518n

Barron, J. J., Ashton, C., & Geary, L. (2006). The Effects of Temperature on pH Measurement. Clare: Rreagecon. Retrieved from http://reagecon.com/pdf/technicalpapers/Effects_of_Temperature_on_pH_v4-_TSP-01-2.pdf

Batra, N., Singh, J., Banerjee, U. C., Patnaik, P. R., & Sobti, R. C. (2002). Production and characterization of a thermostable beta-galactosidase from Bacillus coagulans RCS3. *Biotechnology and Applied Biochemistry*, *36*(Pt 1), 1–6. doi:10.1042/

Bayané, A., & Guiot, S. R. (2010). Animal digestive strategies versus anaerobic digestion bioprocesses for biogas production from lignocellulosic biomass. *Reviews in Environmental Science and Bio/Technology*, *10*(1), 43–62. doi:10.1007/s11157-010-9209-4

Bayer, E. a, Shimon, L. J., Shoham, Y., & Lamed, R. (1998). Cellulosomes-structure and ultrastructure. *Journal of Structural Biology*, *124*(2-3), 221–34.

doi:10.1006/jsbi.1998.4065

BBSRC. (2012). Big score for British biofuel technology. Retrieved from http://www.bbsrc.ac.uk/news/industrial-biotechnology/2012/120127-f-british-biofuel-technology.aspx

BEACON. (2011). A BEACON of light for the green economy. Retrieved from http://www.beaconwales.org/

Beg, Q. K., Bhushan, B., Kapoor, M., & Hoondal, G. S. (2000). Production and characterization of thermostable xylanase and pectinase from Streptomyces sp. QG-11-3. *Journal of Industrial Microbiology and Biotechnology*, *24*(6), 396–402. doi:10.1038/sj.jim.7000010

Behera, S., Arora, R., Nandhagopal, N., & Kumar, S. (2014). Importance of chemical pretreatment for bioconversion of lignocellulosic biomass. *Renewable and Sustainable Energy Reviews*, *36*, 91–106. doi:10.1016/j.rser.2014.04.047

Bentley, S. D., & Parkhill, J. (2004). Comparative genomic structure of prokaryotes. *Annual Review of Genetics*, *38*(13), 771–792. doi:10.1146/annurev.genet.38.072902.094318

Berg, G., Eberl, L., & Hartmann, A. (2005). The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. *Environmental Microbiology*, *7*(11), 1673–85. doi:10.1111/j.1462-2920.2005.00891.x

Bergquist, P. L., Gibbs, M. D., Morris, D. D., Te'o, V. S. J., Saul, D. J., & Morgan, H. W. (1999). Molecular diversity of thermophilic cellulolytic and hemicellulolytic bacteria. *FEMS Microbiology Ecology*, *28*(2), 99–110. doi:10.1016/S0168-6496(98)00078-6

Berlemont, R., & Galleni, M. (2012). Carbohydrate actives enzymes derived from metagenomes: from microbila ecology to enzymology. In R. W. Li (Ed.), *Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies* (1st ed., pp. 185–209). Hauppauge: Nova Science Publishers, Inc.

Bernier, R., Desrochers, M., Jurasek, L., & Paice, M. G. (1983). Isolation and Characterization of a Xylanase from Bacillus subtilis. *Applied and Environmental Microbiology*, *46*(2), 511–514.

Bertoldo, C., & Antranikian, G. (2002). Starch-hydrolyzing enzymes from thermophilic archaea and bacteria. *Current Opinion in Chemical Biology*, *6*, 151–160.

Bertran, M. S., & Dale, B. E. (1985). Enzymatic hydrolysis and recrystallization behavior of initially amorphous cellulose. *Biotechnology and Bioengineering*, *27*(2), 177–81. doi:10.1002/bit.260270212

Beveridge, T. J. (1990). Mechanism of gram variability in select bacteria. *Journal of Bacteriology*, *172*(3), 1609–1620.

Bhadra, B., Rao, R. S., Singh, P. K., Sarkar, P. K., & Shivaji, S. (2008). Yeasts and yeast-like fungi associated with tree bark: diversity and identification of yeasts producing extracellular endoxylanases. *Current Microbiology*, *56*(5), 489–94. doi:10.1007/s00284-008-9108-x

Biocatalysts. (2013). Depol™ 761P - D761P. Parc Nantgarw: Biocatalysts. Retrieved from www.biocatalysts.com

Biocatalysts Ltd. (2015a). Depol™ 333MDP-D333MDP. Cardiff: www.biocatalysts.com.

Biocatalysts Ltd. (2015b). Depol™ 686L-D686L. Cardiff: www.biocatalysts.com.

Bioneer Corp. (2014). ExiProgen EC1 protein synthesis kit. Daejeon: bioneer.com.

Birur, D. K., Hertel, T. W., & Tyner, W. E. (2008). *Impact of Biofuel Production on World Agricultural Markets: A Computable General Equilibrium Analysis*.

Bischoff, K. M., Rooney, A. P., Li, X.-L., Liu, S., & Hughes, S. R. (2006). Purification and characterization of a family 5 endoglucanase from a moderately thermophilic strain of Bacillus licheniformis. *Biotechnology Letters*, *28*(21), 1761–5. doi:10.1007/s10529-006-9153-0

Blake, A. W., McCartney, L., Flint, J. E., Bolam, D. N., Boraston, A. B., Gilbert, H. J., & Knox, J. P. (2006). Understanding the biological rationale for the diversity of cellulose-directed carbohydrate-binding modules in prokaryotic enzymes. *The Journal of Biological Chemistry*, *281*(39), 29321–9. doi:10.1074/jbc.M605903200

Blot, N., Mavathur, R., Geertz, M., Travers, A., & Muskhelishvili, G. (2006). Homeostatic regulation of supercoiling sensitivity coordinates transcription of the bacterial genome. *EMBO Reports*, *7*(7), 710–715. doi:10.1038/sj.embor.7400729

Bodilis, J., Nsigue-Meilo, S., Besaury, L., & Quillet, L. (2012). Variable copy number, intra-genomic heterogeneities and lateral transfers of the 16S rRNA gene in Pseudomonas. *PloS One*, *7*(4), e35647. doi:10.1371/journal.pone.0035647

Boehm, H. J., Boehringer, M., Bur, D., Gmuender, H., Huber, W., Klaus, W., … Mueller, F. (2000). Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization. A promising alternative to random screening. *Journal of Medicinal Chemistry*, *43*(14), 2664–2674. doi:10.1021/jm000017s

Boer, W. De, Folman, L. B., Summerbell, R. C., & Boddy, L. (2005). Living in a fungal world: impact of fungi on soil bacterial niche development. *FEMS Microbiology Reviews*, *29*(4), 795–811. doi:10.1016/j.femsre.2004.11.005

Bondos, S. E., & Bicknell, A. (2003). Detection and prevention of protein aggregation before, during, and after purification. *Analytical Biochemistry*, *316*(2), 223–231. doi:10.1016/S0003-2697(03)00059-9

Bonnette, R. E. (2013). *Food Biotechnology Consultation Note to the File BNF No . 000095*. Silver Spring. Retrieved from www.fda.gov/Food/FoodScienceResearch/Biotechnology/Submissions/ucm155608.htm

Borglum, G. B. (1980). *Starch Hydrolysis for Ethanol Production*. Elkhart, Indiana.

BP plc. (2013). Biofuels operations & technologies. Retrieved from http://www.bp.com/sectiongenericarticle.do?categoryId=9030041&contentId=7055175

Bradshaw, D. J., & Marsh, P. D. (1998). Analysis of pH-driven disruption of oral microbial communities in vitro. *Caries Research*, *32*(6), 456–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9745120

Bragger, J. M., Daniel, R. M., Coolbear, T., & Morgan, H. W. (1989). Very stable enzymes from extremely thermophilic archaebacteria and eubacteria. *Applied Microbiology and Biotechnology*, *31*(5-6), 556–561. doi:10.1007/BF00270794

Brandt, A., Gräsvik, J., Hallett, J. P., & Welton, T. (2013). Deconstruction of lignocellulosic biomass with ionic liquids. *Green Chemistry*, *15*(3), 550. doi:10.1039/c2gc36364j

Brandt, A., Ray, M. J., To, T. Q., Leak, D. J., Murphy, R. J., & Welton, T. (2011). Ionic liquid pretreatment of lignocellulosic biomass with ionic liquid–water mixtures. *Green Chemistry*, *13*(9), 2489. doi:10.1039/c1gc15374a

Brechtel, E., Matuschek, M., Hellberg, A., Egelseer, E. M., Schmid, R., & Bahl, H. (1999). Cell wall of Thermoanaerobacterium thermosulfurigenes EM1: isolation of its components and attachment of the xylanase XynA. *Archives of Microbiology*, *171*(3), 159–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10201095

Bresciani, F. R., Santi, L., Macedo, A. J., Abraham, W.-R., Vainstein, M. H., & Beys-da-Silva, W. O. (2013). Production and activity of extracellular lipase from Luteibacter sp. *Annals of Microbiology*, *64*(1), 251–258. doi:10.1007/s13213-013-0657-2

Brodeur, G., Yau, E., Badal, K., Collier, J., Ramachandran, K. B., & Ramakrishnan, S. (2011). Chemical and physicochemical pretreatment of lignocellulosic biomass: a review. *Enzyme Research*, *2011*, 787532. doi:10.4061/2011/787532

Bronnenmeier, K., Meissner, H., Stocker, S., & Staudenbauer, W. L. (1995). a-D-Glucuronidases from the xylanolytic thermophiles Clostridium stercorarium and Thermoanaerobacterium saccharolyticum. *Microbiology*, *141*, 2033–2040.

Brookes, A. J. (1999). The essence of SNPs. *Gene*, *234*(2), 177–86. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10395891

Brosius, J., Palmer, M. L., Kennedy, P. J., & Noller, H. F. (1978). Complete nucleotide sequence of a 16S ribosomal RNA gene from Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, *75*(10), 4801–5. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=336208&tool=pmcentrez&rendertype=abstract

Bryant, D. N., Morris, S. M., Leemans, D., Fish, S. a, Taylor, S., Carvell, J., … Gallagher, J. a. (2011). Modelling real-time simultaneous saccharification and fermentation of lignocellulosic biomass and organic acid accumulation using dielectric spectroscopy. *Bioresource Technology*, *102*(20), 9675–82. doi:10.1016/j.biortech.2011.07.084

Bulkeley, H., & Newell, P. (2015). *Governing Climate Change*. (T. G. Weiss, Ed.) (2nd ed.). Abingdon: Routledge.

Burhan, A., Nisa, U., Gökhan, C., Ömer, C., Ashabil, A., & Osman, G. (2003). Enzymatic

properties of a novel thermostable, thermophilic, alkaline and chelator resistant amylase from an alkaliphilic Bacillus sp. isolate ANT-6. *Process Biochemistry*, *38*(10), 1397–1403. doi:10.1016/S0032-9592(03)00037-2

Bussemaker, M. J., & Zhang, D. (2013). Effect of Ultrasound on Lignocellulosic Biomass as a Pretreatment for Biorefinery and Biofuel Applications. *Industrial & Engineering Chemistry Research*, *52*, 3563–3580. doi:10.1021/ie3022785

Caccavo, F., Lonergan, D. J., Lovley, D. R., Davis, M., Stolz, J. F., & Mcinerneyl, M. J. (1994). Oxidizing Dissimilatory Metal-Reducing Microorganism. *Microbiology*, *60*(10), 3752–3759.

Cai, T., Qian, L., Cai, S., & Chen, L. (2011). Biodegradation of benazolin-ethyl by strain Methyloversatilis sp. cd-1 isolated from activated sludge. *Current Microbiology*, *62*(2), 570–577. doi:10.1007/s00284-010-9746-7

Cairns, A., Gallagher, J., Hatch, R., & Humphreys, M. (2007). *A future for UK grassland in energy production?*

Callaghan, A. V., Wawrik, B., Ní Chadhain, S. M., Young, L. Y., & Zylstra, G. J. (2008). Anaerobic alkane-degrading strain AK-01 contains two alkylsuccinate synthase genes. *Biochemical and Biophysical Research Communications*, *366*(1), 142–148. doi:10.1016/j.bbrc.2007.11.094

Carlson, E. D., Gan, R., Hodgman, C. E., & Jewett, M. C. (2012). Cell-free protein synthesis: Applications come of age. *Biotechnology Advances*, *30*(5), 1185–1194. doi:10.1016/j.biotechadv.2011.09.016

Carpita, N. C. (1996). Structure and Biogenesis of the Cell Walls of Grasses. *Annual Review of Plant Physiology and Plant Molecular Biology*, *47*, 445–476. doi:10.1146/annurev.arplant.47.1.445

CAZy. (2012). Glycoside hydrolase family classification. Retrieved from http://www.cazy.org/Glycoside-Hydrolases.html

Center for Microbial Ecology. (2012). Seq Match Help - Ribosomal Database Project Wiki. Retrieved July 9, 2015, from http://rdp.cme.msu.edu/wiki/index.php/Seq_Match_Help

Chakraborty, S., Khopade, A., Biao, R., Jian, W., Liu, X.-Y., Mahadik, K., … Kokare, C. (2011). Characterization and stability studies on surfactant, detergent and oxidant stable α-amylase from marine haloalkaliphilic Saccharopolyspora sp. A9. *Journal of Molecular Catalysis B: Enzymatic*, *68*(1), 52–58. doi:10.1016/j.molcatb.2010.09.009

Chakravorty, S., Helb, D., Burday, M., Connell, N., & Alland, D. (2007). A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *Journal of Microbial Methods*, *69*(2), 330–339. doi:10.1016/j.mimet.2007.02.005.A

Chandrashekharaiah, K. S., Murthy, V. K., Narayanaswamy, M., Murthy, S., & Swamy, N. R. (2013). CHARACTERIZATION OF ALPHA-AMYLASE FROM THE SEEDS OF Mucuna pruriens. *Journal of Experimental Biology and Agricultural Sciences*, *1*(5),

367–374.

Charlton, A., Elias, R., Fish, S., Fowler, P., & Gallagher, J. (2009). The biorefining opportunities in Wales: Understanding the scope for building a sustainable, biorenewable economy using plant biomass. *Chemical Engineering Research and Design*, *87*(9), 1147–1161. doi:10.1016/j.cherd.2009.06.013

Chen, Y., Cheng, J. J., & Creamer, K. S. (2008). Inhibition of anaerobic digestion process: A review. *Bioresource Technology*, *99*(10), 4044–4064. doi:10.1016/j.biortech.2007.01.057

Cheng, H., & Wang, L. (2013). Lignocelluloses Feedstock Biorefinery as Petrorefinery Substitutes. In M. D. Matovic (Ed.), *Biomass Now - Sustainable Growth and Use* (1st ed., pp. 347–388). Rijeka: InTech. doi:10.5772/51491

Chethana, S., Pratap, B., Roy, S., Jaiswal, A., Shruthi, S., & Vedamurthy, A. (2011). Bioethanol production from rice water waste: a low cost motor fuel. *Pharmacologyonline*, *3*, 125–134.

Chimata, M. K., Chetty, C. S., & Suresh, C. (2011). Fermentative production and thermostability characterization of α amylase from Aspergillus species and its application potential evaluation in desizing of cotton cloth. *Biotechnology Research International*, *2011*, 323891. doi:10.4061/2011/323891

Chin, K., & H'ng, P. (2013). A Real Story of Bioethanol from Biomass: Malaysia Perspective. In M. D. Matovic (Ed.), *Biomass Now - Sustainable Growth and Use* (1st ed., pp. 329–346). Rijeka: InTech. doi:10.5772/54528

Cho, H.-Y., Kim, Y.-W., Kim, T.-J., Lee, H.-S., Kim, D.-Y., Kim, J.-W., … Park, K.-H. (2000). Molecular characterization of a dimeric intracellular maltogenic amylase of Bacillus subtilis SUH4-2. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, *1478*(2), 333–340. doi:10.1016/S0167-4838(00)00037-6

Cieslewicz, M. J., Kasper, D. L., Wang, Y., & Wessels, M. R. (2001). Functional analysis in type Ia group B Streptococcus of a cluster of genes involved in extracellular polysaccharide production by diverse species of streptococci. *Journal of Biological Chemistry*, *276*(1), 139–146. doi:10.1074/jbc.M005702200

Claeyssens, M., & Henrissat, B. (1992). Specificity mapping of cellulolytic enzymes: classification into families of structurally related proteins confirmed by biochemical analysis. *Protein Science : A Publication of the Protein Society*, *1*(10), 1293–1297. doi:10.1002/pro.5560011008

Cock, P. J. a, & Whitworth, D. E. (2010). Evolution of relative reading frame bias in unidirectional prokaryotic gene overlaps. *Molecular Biology and Evolution*, *27*(4), 753–756. doi:10.1093/molbev/msp302

Coenye, T., Laevens, S., Willems, A., Ohle, M., Hannant, W., Govan, J. R. W., … Differential, T. (2001). Burkholderia caledonica sp. nov., two new species isolated from the environment, animals and human clinical samples. *International Journal of Systematic and Evolutionary Microbiology*, *51*, 1099–1107.

Coil, D., Jospin, G., & Darling, A. E. (2014). Genome analysis A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Oxford Journals*, *2*, 1–3. doi:10.1093/bioinformatics/btu661

Colin, V. L., Rodríguez, A., & Cristóbal, H. A. (2011). The role of synthetic biology in the design of microbial cell factories for biofuel production. *Journal of Biomedicine & Biotechnology*, *2011*, 601834. doi:10.1155/2011/601834

Collins, T., Gerday, C., & Feller, G. (2005). Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiology Reviews*, *29*(1), 3–23. doi:10.1016/j.femsre.2004.06.005

Connor, M. R., & Atsumi, S. (2010). Synthetic biology guides biofuel production. *Journal of Biomedicine & Biotechnology*, *2010*. doi:10.1155/2010/541698

Coombs, J. (1996). *Bioconversion Assessment Study* (First.). Luxembourg: European Commission.

Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nature Reviews. Molecular Cell Biology*, *6*(11), 850–61. doi:10.1038/nrm1746

Cotula, L., Dyer, N., & Vermeulen, S. (2008). *Fuelling exclusion? The biofuels boom and poor people's access to land*. London.

Couturier, M., Navarro, D., Olivé, C., Chevret, D., Haon, M., Favel, A., … Berrin, J.-G. (2012). Post-genomic analyses of fungal lignocellulosic biomass degradation reveal the unexpected potential of the plant pathogen Ustilago maydis. *BMC Genomics*, *13*(1), 57. doi:10.1186/1471-2164-13-57

Crittenden, R., Karppinen, S., Ojanen, S., Tenkanen, M., Fagerstrm, R., Mtt, J., … Poutanen, K. (2002). In vitro fermentation of cereal dietary fibre carbohydrates by probiotic and intestinal bacteria. *Journal of the Science of Food and Agriculture*, *82*(8), 781–789. doi:10.1002/jsfa.1095

Crowhurst, L., Mawdsley, P. R., Perez-Arlandis, J. M., Salter, P. a., & Welton, T. (2003). Solvent - solute interactions in ionic liquids. *Physical Chemistry Chemical Physics*, *5*(13), 2790. doi:10.1039/b303095d

Crowther, M. (1956). United States Patent ice. USA: United States Patent Office.

Cruz, A. G., Scullin, C., Mu, C., Cheng, G., Stavila, V., Varanasi, P., … Singh, S. (2013). Impact of high biomass loading on ionic liquid pretreatment. *Biotechnology for Biofuels*, *6*, 52–61.

da Costa Lopes, A. M., João, K. G., Rubik, D. F., Bogel-Łukasik, E., Duarte, L. C., Andreaus, J., & Bogel-Łukasik, R. (2013). Pre-treatment of lignocellulosic biomass using ionic liquids: wheat straw fractionation. *Bioresource Technology*, *142*, 198–208. doi:10.1016/j.biortech.2013.05.032

Damiano, V. B., Ward, R., Gomes, E., Alves-Prado, H. F., & Da Silva, R. (2006). Purification and characterization of two xylanases from alkalophilic and thermophilic Bacillus licheniformis 77-2. *Applied Biochemistry and Biotechnology*, *129-132*(1), 289–302. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21729216

Danielsen, F., Beukema, H., Burgess, N. D., Parish, F., Brühl, C. a, Donald, P. F., … Fitzherbert, E. B. (2009). Biofuel plantations on forested lands: double jeopardy for biodiversity and climate. *Conservation Biology : The Journal of the Society for Conservation Biology*, *23*(2), 348–58. doi:10.1111/j.1523-1739.2008.01096.x

Darling, A. C. E., Mau, B., Blattner, F. R., & Perna, N. T. (2004). Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements. *Genome Research*, *14*, 1394–1403. doi:10.1101/gr.2289704

David, C., Fornasier, R., Greindl-Fallon, C., & Vanlautem, N. (1985). Enzymatic hydrolysis and bacterian hydrolysis - fermentation of eucalyptus wood pretreated with sodium hypochlorite. *Biotechnology and Bioengineering*, *27*(11), 1591–1595. doi:10.1002/bit.260271112

Davies, G., & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure*, *3*(9), 853–859. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8535779

Dawson, L., & Boopathy, R. (2008). Cellulosic ethanol production from sugarcane bagasse without enzymatic saccharification. *BioResources*, *3*(2), 452–460.

Dawson, R. M. C. (2002). *Data for biochemical research*. (D. C. Elliott, W. H. Elliott, & K. M. Jones, Eds.) (3rd ed.). Oxford: Oxford University Press.

de Boer, A., de Boer, A., Priest, F., Priest, F., Diderichsen, B., & Diderichsen, B. (1994). On the Industrial Use of Bacillus licheniformis: a Review. *Applied Microbiology and Biotechnology*, *40*(5), 595–598. Retrieved from http://apps.isiknowledge.com/full_record.do?product=UA&search_mode=Refine&qid=3&SID=Z15A4ee@O7e8IG7hL75&page=1&doc=1&colname=WOS

de Frias, J. A., & Feng, H. (2013). Switchable butadiene sulfone pretreatment of Miscanthus in the presence of water. *Green Chemistry*, *15*, 1067–1078. doi:10.1039/c3gc37099b

de Lucena, B. T. L., Silva, G. G. Z., dos Santos, B. M., Dias, G. M., Amaral, G. R. S., Moreira, A. P. B., … Thompson, F. L. (2012). Genome sequences of the ethanol-tolerant lactobacillus vini strains LMG 23202T and JP7.8.9. *Journal of Bacteriology*, *194*(11), 3018. doi:10.1128/JB.00446-12

De Vecchi, E., & Drago, L. (2006). Lactobacillus sporogenes or Bacillus coagulans: misidentification or mislabelling? *International Journal of Probiotics and Prebiotics*, *1*(1), 3–10.

de Vrije, T., Bakker, R. R., Budde, M. A., Lai, M. H., Mars, A. E., & Claassen, P. A. (2009). Efficient hydrogen production from the lignocellulosic energy crop Miscanthus by the extreme thermophilic bacteria Caldicellulosiruptor saccharolyticus and Thermotoga neapolitana. *Biotechnology for Biofuels*, *2*(1), 12. doi:10.1186/1754-6834-2-12

Declerck, N., Machius, M., Joyet, P., Wiegand, G., Huber, R., & Gaillardin, C. (2003). Hyperthermostabilization of Bacillus licheniformis -amylase and modulation of its stability over a 50 C temperature range. *Protein Engineering Design and Selection*, *16*(4), 287–293. doi:10.1093/proeng/gzg032

Declerck, N., Machius, M., Wiegand, G., Huber, R., & Gaillardin, C. (2000). Probing structural determinants specifying high thermostability in Bacillus licheniformis alpha-amylase. *Journal of Molecular Biology*, *301*(4), 1041–57. doi:10.1006/jmbi.2000.4025

Dee, S., & Bell, A. T. (2011). Effects of reaction conditions on the acid-catalyzed hydrolysis of miscanthus dissolved in an ionic liquid. *Green Chemistry*, *13*(6), 1467. doi:10.1039/c1gc15317j

Department of Energy & Climate Change. (2012). *UK Bioenergy Strategy*. London.

Dodd, D., & Cann, I. K. (2009). Enzymatic deconstruction of xylan for biofuel production. *GCB Bioenergy*, *1*(1), 2–17. doi:10.1111/j.1757-1707.2009.01004.x.Enzymatic

Dong, G., Vieille, C., Savchenko, A., & Zeikus, J. G. (1997). Cloning, sequencing, and expression of the gene encoding extracellular alpha-amylase from Pyrococcus furiosus and biochemical characterization of the recombinant enzyme. *Applied and Environmental Microbiology*, *63*(9), 3569–3576.

Donkersloot, J. a, de Leon, H. a, Chassy, B. M., & Krichevsky, M. I. (1976). Analysis of the exudate produced by Streptococcus mutans SL-1 colonies of sucrose-containing agar media. *Applied and Environmental Microbiology*, *32*(3), 448–450.

Donovan, K. (2013). Iogen Bio-Products aquired by Novozymes. Ottawa: Iogen Corporation.

Doronina, N. V., Kaparullina, E. N., & Trotsenko, Y. A. (2014). Methyloversatilis thermotolerans sp. nov., a novel thermotolerant facultative methylotroph isolated from a hot spring. *International Journal of Systematic and Evolutionary Microbiology*, *64*(PART 1), 158–164. doi:10.1099/ijs.0.055046-0

Drapcho, C. M., Nhuan, N. P., & Walker, T. H. (2008). *Biofuels Engineering Process Technology* (1st ed.). New York: McGraw Hill. doi:10.1036/0071487492

Drax Group plc. (2013). Our business Drax. Retrieved from http://www.drax.com/aboutus/ourbusiness/

Du Toit, A. (2015). Bacterial pathogenesis: Copycat Burkholderia make a move. *Nature Reviews Microbiology*, *13*(6), 330–330. doi:10.1038/nrmicro3488

Dumon, C., Song, L., Bozonnet, S., Fauré, R., & O'Donohue, M. J. (2012). Progress and future prospects for pentose-specific biocatalysts in biorefining. *Process Biochemistry*, *47*(3), 346–357. doi:10.1016/j.procbio.2011.06.017

Eckert, B., Weber, O. B., Kirchhof, G., Halbritter, a, Stoffels, M., & Hartmann, a. (2001). Azospirillum doebereinerae sp. nov., a nitrogen-fixing bacterium associated with the C4-grass Miscanthus. *International Journal of Systematic and Evolutionary Microbiology*, *51*(Pt 1), 17–26. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11211255

Eiland, F., Klamer, M., Lind, a.-M., Leth, M., & Bååth, E. (2001). Influence of Initial C/N Ratio on Chemical and Microbial Composition during Long Term Composting of Straw. *Microbial Ecology*, *41*(3), 272–280. doi:10.1007/s002480000071

Elsayed, S., & Zhang, K. (2005). Bacteremia Caused by Clostridium intestinale Bacteremia Caused by Clostridium intestinale, *43*(4), 11–14. doi:10.1128/JCM.43.4.2018

EMBL-EBI. (2014). ClustalW2 Multiple Sequence Alignment. Retrieved March 6, 2015, from http://www.ebi.ac.uk/Tools/msa/clustalw2/

Encylopaedia Britannica Inc. (2014). Bacillus. Retrieved from http://www.britannica.com/EBchecked/topic/47965/bacillus

Enogen Advisory Council. (2011). Enogen /Corn Amylase/Event 3272 ™ Product Overview. In *Advisory Council Meething* (pp. 1–18). Chicago: Syngenta. Retrieved from http://www.syngenta.com/country/us/en/agriculture/seeds/corn/enogen/stewardship/Documents/June 14th, 2011/Enogen Overview.pdf

EPOBIO. (2006). *Cell Wall Saccharification*. (R. Moller, Ed.) (1st ed.). York: CPL Press Science Publishers.

Epting, K. L., Vieille, C., Zeikus, J. G., & Kelly, R. M. (2005). Influence of divalent cations on the structural thermostability and thermal inactivation kinetics of class II xylose isomerases. *FEBS Journal*, *272*(6), 1454–1464. doi:10.1111/j.1742-4658.2005.04577.x

Esposito, D., & Antonietti, M. (2015). Redefining biorefinery: the search for unconventional building blocks for materials. *Chem. Soc. Rev.* doi:10.1039/C4CS00368C

Ezeji, T. C., Qureshi, N., & Blaschek, H. P. (2007). Bioproduction of butanol from biomass: from genes to bioreactors. *Current Opinion in Biotechnology*, *18*(3), 220–7. doi:10.1016/j.copbio.2007.04.002

Fang, H. H. P., & Liu, H. (2002). Effect of pH on hydrogen production from glucose by a mixed culture. *Bioresource Technology*, *82*(1), 87–93. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11858207

Faulon, J.-L., Carlson, G. A., & Hatcher, P. G. (1994). A three-dimensional model for lignocellulose from gymnospermous wood. *Organic Geochemistry*, *21*(12), 1169–1179. doi:10.1016/0146-6380(94)90161-9

Feller, G., Lonhienne, T., Deroanne, C., Libioulle, C., Beeumenq, V., & Gerday, C. (1992). Purification, characterisation, and nucleotide sequence of the thermolabile a-amylase from the Antartic psychrotroph Alteromonas haloplanctis A23. *The Journal of Biological Chemistry*, *267*, 5217–5221.

Felsenstein, J. (1985). Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Society for the Study of Evolution*, *39*(4), 783–791.

Ferkovich, S., & Mayer, M. (1975). Localization and specifity of pheromone degrading enzyme(s) from antennae of Trichoplusia ni. In D. A. Denton & J. P. Coghlan (Eds.), *Olfaction and taste V* (1st ed., pp. 337–342). London: Academic Press, Inc.

Fernandes, S., & Murray, P. (2010). Metabolic engineering for improved microbial pentose fermentation. *Bioengineered Bugs*, *1*(6), 424–428.

doi:10.4161/bbug.1.6.12724

Ferrer, M., Golyshina, O., Beloqui, A., & Golyshin, P. N. (2007). Mining enzymes from extreme environments. *Current Opinion in Microbiology*, *10*(3), 207–14. doi:10.1016/j.mib.2007.05.004

FitzPatrick, M., Champagne, P., Cunningham, M. F., & Whitney, R. a. (2010). A biorefinery processing perspective: treatment of lignocellulosic materials for the production of value-added products. *Bioresource Technology*, *101*(23), 8915–22. doi:10.1016/j.biortech.2010.06.125

Fontes, C. M. G. a, & Gilbert, H. J. (2010). Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Annual Review of Biochemistry*, *79*, 655–681. doi:10.1146/annurev-biochem-091208-085603

Freier, D., Mothershed, C. P., & Wiegel, J. (1988). Characterization of Clostridium thermocellum JW20. *Applied and Environmental Microbiology*, *54*(1), 204–211. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=202422&tool=pmcentrez&rendertype=abstract

French, C. E. (2009). Synthetic biology and biomass conversion: a match made in heaven? *Journal of the Royal Society, Interface / the Royal Society*, *6 Suppl 4*(May), S547–58. doi:10.1098/rsif.2008.0527.focus

Freundlieb, S., & Boos, W. (1986). a-amylase of Escherichia coli, mapping and cloning of the structural gene, mal & and identification of its product as a periplasmic protein. *The Journal of Biological Chemistry*, *261*(6), 2946–2953.

Fujii, M., Murakami, S., Yamada, Y., Ona, T., & Nakamura, T. (1981). A Kinetic Equation for Hydrolysis of Polysaccharides by Mixed Exo- and Endoenzyme Systems. *Biotechnology and Bioengineering*, *23*(6), 1393–1398.

Gao, D., Uppugundla, N., Chundawat, S. P., Yu, X., Hermanson, S., Gowda, K., … Dale, B. E. (2011). Hemicellulases and auxiliary enzymes for improved conversion of lignocellulosic biomass to monosaccharides. *Biotechnology for Biofuels*, *4*(1), 50–61. doi:10.1186/1754-6834-4-5

Garimella, R., Halye, J. L., Harrison, W., Klebba, P. E., & Rice, C. V. (2009). Conformation of the phosphate D-alanine zwitterion in bacterial teichoic acid from nuclear magnetic resonance spectroscopy. *Biochemistry*, *48*(39), 9242–9249. doi:10.1021/bi900503k

Gavrilescu, M., & Chisti, Y. (2005). Biotechnology-a sustainable alternative for chemical industry. *Biotechnology Advances*, *23*(7-8), 471–99. doi:10.1016/j.biotechadv.2005.03.004

Gellert, M., O'Dea, M. H., Itoh, T., & Tomizawa, J. (1976). Novobiocin and coumermycin inhibit DNA supercoiling catalyzed by DNA gyrase. *Proceedings of the National Academy of Sciences of the United States of America*, *73*(12), 4474–4478. doi:10.1073/pnas.73.12.4474

George, S. P., Ahmad, a, & Rao, M. B. (2001). A novel thermostable xylanase from

Thermomonospora sp.: influence of additives on thermostability. *Bioresource Technology*, *78*(3), 221–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11341679

Gerardi, M. H. (2003). *The Microbiology of Anaerobic Digesters*. (M. H. Gerardi, Ed.) (1st ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Ghatora, S. K., Chadha, B. S., Badhan, A. K., Saini, H. S., & Bhat, M. K. (2006). IDENTIFICATION AND CHARACTERIZATION OF DIVERSE XYLANASES FROM THERMOPHILIC AND THERMOTOLERANT Fungi. *BioResources*, *1*(2006), 18–33.

Gilbert, H. J., & Hazlewood, G. P. (1993). Bacterial cellulases and xylanases. *Journal of General Microbiology*, *139*, 187–194.

Gírio, F. M., Fonseca, C., Carvalheiro, F., Duarte, L. C., Marques, S., & Bogel-Łukasik, R. (2010). Hemicelluloses for fuel ethanol: A review. *Bioresource Technology*, *101*(13), 4775–800. doi:10.1016/j.biortech.2010.01.088

Glick, B. R. (2012). Plant growth-promoting bacteria: mechanisms and applications. *Scientifica*, *2012*, 963401. doi:10.6064/2012/963401

GMO Compass. (2010). Xylanase - GMO Database. Retrieved from http://www.gmo-compass.org/eng/database/enzymes/96.xylanase.html

Goldstein, E. J. C., Citron, D. M., Peraino, V. a, & Cross, S. a. (2003). Desulfovibrio desulfuricans Bacteremia and Review of Human Desulfovibrio Infections Desulfovibrio desulfuricans Bacteremia and Review of Human Desulfovibrio Infections, *41*(6), 2752–2754. doi:10.1128/JCM.41.6.2752

Golyshina, O. V, Golyshin, P. N., Timmis, K. N., & Ferrer, M. (2006). The “pH optimum anomaly” of intracellular enzymes of Ferroplasma acidiphilum. *Environmental Microbiology*, *8*(3), 416–425. doi:10.1111/j.1462-2920.2005.00907.x

Gomes, I., Gomes, J., & Steiner, W. (2003). Highly thermostable amylase and pullulanase of the extreme thermophilic eubacterium Rhodothermus marinus: production and partial characterization. *Bioresource Technology*, *90*(2), 207–214. doi:10.1016/S0960-8524(03)00110-X

González-Fernández, R., Prats, E., & Jorrín-Novo, J. V. (2010). Proteomics of plant pathogenic fungi. *Journal of Biomedicine & Biotechnology*, *2010*, 932527. doi:10.1155/2010/932527

Gool, M. P. Van. (2012). *Targeted discovery and functional characterisation of complex-xylan degrading enzymes*. Wageningen University.

Goto, Y., Calciano, L. J., & Fink, A. L. (1990). Acid-induced folding of proteins. *PNAS*, *87*(January), 573–577.

Govan, J. R., & Deretic, V. (1996). Microbial Pathogenesis in Cystic Fibrosis: Mucoid Pseudomonas aeruginosa and Burkholderia cepacia. *Microbiology and Molecular Biology Reviews*, *60*(3), 539–574.

Graenacher, C. (1934). Cellulose solution. Switzerland: United States Patent Office.

Gray, G. L., Mainzer, S. E., Rey, M. W., Lamsa, M. H., Kindle, K. L., Carmona, C., &

Requadt, C. (1986). Structural Genes Encoding the Thermophilic a-Amylases of Bacillus stearothermophilus and Bacillus licheniformis. *Journal of Bacteriology*, *166*(2), 635–643.

Gregersen, T. (1978). Rapid method for the distinction of Gram-negative from Gram-positive bacteria. *Applied Microbiology and Biotechnology*, *5*, 123–127.

Grewell, D. (2008). Application of ultrasound to pretreat corn or biomass for ethanol production. Iowa: Iowa State University. Retrieved from http://dgrewell.public.iastate.edu/research/biofuels/ultrasound_pretreatment.pdf

Gu, Y., Ding, Y., Ren, C., Sun, Z., Rodionov, D. a, Zhang, W., … Jiang, W. (2010). Reconstruction of xylose utilization pathway and regulons in Firmicutes. *BMC Genomics*, *11*(Figure 1), 255. doi:10.1186/1471-2164-11-255

Guevara, C., & Zambrano, M. M. (2006). Sugarcane cellulose utilization by a defined microbial consortium. *FEMS Microbiology lLtters*, *255*(1), 52–8. doi:10.1111/j.1574-6968.2005.00050.x

Guglielmetti, S., Basilico, R., Taverniti, V., Arioli, S., Piagnani, C., & Bernacchi, A. (2013). Luteibacter rhizovicinus MIMR1 promotes root development in barley (Hordeum vulgare L.) under laboratory conditions. *World Journal of Microbiology & Biotechnology*, *29*(11), 2025–32. doi:10.1007/s11274-013-1365-6

Guipaud, O., Marguet, E., Noll, K. M., de la Tour, C. B., & Forterre, P. (1997). Both DNA gyrase and reverse gyrase are present in the hyperthermophilic bacterium Thermotoga maritima. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(20), 10606–10611. doi:10.1073/pnas.94.20.10606

Guo, G., Liu, Z., Xu, J., Liu, J., Dai, X., Xie, D., … Fu, Y. (2012). Purification and characterization of a xylanase from Bacillus subtilis isolated from the degumming line. *Journal of Basic Microbiology*, *52*(4), 419–428. doi:10.1002/jobm.201100262

Guo, T., Tang, Y., Zhang, Q. Y., Du, T. F., Liang, D. F., Jiang, M., & Ouyang, P. K. (2012). Clostridium beijerinckii mutant with high inhibitor tolerance obtained by low-energy ion implantation. *Journal of Industrial Microbiology and Biotechnology*, *39*(3), 401–407. doi:10.1007/s10295-011-1017-5

Gupta, P. K. (2014). Chemistry of the Gene: 2. Synthesis, Modifications and Repair of DNA. In *Genetics* (4th ed., pp. 297–323). New Delhi: Rastogi Publications.

Gupta, P., Samant, K., & Sahu, A. (2012). Isolation of cellulose-degrading bacteria and determination of their cellulolytic potential. *International Journal of Microbiology*, *2012*, 578925. doi:10.1155/2012/578925

Gupta, R., Gigras, P., Mohapatra, H., Goswami, V. K., & Chauhan, B. (2003). Microbial α-amylases: a biotechnological perspective. *Process Biochemistry*, *38*(11), 1599–1616. doi:10.1016/S0032-9592(03)00053-0

Gutierrez-Nava, A., Herrera-Herrera, A., Mayorga-Reyes, L., Salgado, L. M., & Ponce-Noyola, T. (2003). Expression and Characterization of the celcflB Gene from Cellulomonas flavigena Encoding an Endo-b-1,4-Glucanase. *Current Microbiology*,

47(5), 359–363. doi:10.1007/s00284-002-4016-y

Halebian, S., Harris, B., Finegold, S. M., & Rolfe, R. D. (1981). Rapid method that aids in distinguishing gram-positive from gram-negative anaerobic bacteria. *Journal of Clinical Microbiology*, *13*(3), 444–448.

Halford, N. (2011). The Role of Plant Breeding and Biotechnology in Meeting the Challenge of Global Warming. In E. Carayannis (Ed.), *Planet Earth 2011: Global Warming Challenges and Opportunities for Policy and Practice* (First., pp. 493–507). Rijeka: InTech. Retrieved from http://www.intechopen.com/books/planet-earth-2011-global-warming-challenges-and-opportunities-for-policy-and-practice/the-role-of-plant-breeding-and-biotechnology-in-meeting-the-challenge-of-global-warming

Hall, J. R., Mitchell, K. R., Jackson-Weaver, O., Kooser, A. S., Cron, B. R., Crossey, L. J., & Takacs-Vesbach, C. D. (2008). Molecular characterization of the diversity and distribution of a thermal spring microbial community by using rRNA and metabolic genes. *Applied and Environmental Microbiology*, *74*(15), 4910–22. doi:10.1128/AEM.00233-08

Hall, R. M. (2002). Chapter 2 - Gene cassettes and integrons: moving single genes. In M. Syvanen & I. K. Clarence (Eds.), *Horizontal Gene Transfer* (Second., pp. 19–28). London: Academic Press. doi:http://dx.doi.org/10.1016/B978-012680126-2/50005-0

Hallac, B. B., Sannigrahi, P., Pu, Y., Ray, M., Murphy, R. J., & Ragauskas, A. J. (2010). Effect of Ethanol Organosolv Pretreatment on Enzymatic Hydrolysis of Buddleja daWidii Stem Biomass, 1467–1472.

Hallett, J. P., & Welton, T. (2011). Room-temperature ionic liquids: solvents for synthesis and catalysis. 2. *Chemical Reviews*, *111*(5), 3508–76. doi:10.1021/cr1003248

Halsall, D. M., & Gibson, a H. (1985). Cellulose Decomposition and Associated Nitrogen Fixation by Mixed Cultures of Cellulomonas gelida and Azospirillum Species or Bacillus macerans. *Applied and Enviromental Microbiology*, *50*(4), 1021–1026. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=291786&tool=pmcentrez&rendertype=abstract

Hameed, P. S., & Paulpandian, A. L. (1987). Quantitative analysis of carbohydrases in the crystalline style of some intertidal bivalve molluscs. *Proceedings: Animal Sciences*, *96*(1), 41–47. doi:10.1007/BF03179390

Han, D., Xu, H., Puranik, R., & Xu, Z. (2014). Natural transformation of Thermotoga sp. strain RQ7. *BMC Biotechnology*, *14*(1), 39–48. doi:10.1186/1472-6750-14-39

Han, Y., Agarwal, V., Dodd, D., Kim, J., Bae, B., Mackie, R. I., … Cann, I. K. O. (2012). Biochemical and structural insights into xylan utilization by the thermophilic bacterium Caldanaerobius polysaccharolyticus. *The Journal of Biological Chemistry*, *287*(42), 34946–60. doi:10.1074/jbc.M112.391532

Hanshew, A. S., Mason, C. J., Raffa, K. F., & Currie, C. R. (2013). Minimization of

chloroplast contamination in 16S rRNA gene pyrosequencing of insect herbivore bacterial communities. *Journal of Microbiological Methods*, *95*(2), 149–155. doi:10.1016/j.mimet.2013.08.007

Harish, B. ., Janaki Ramaiah, M., & Babu Uppuluri, K. (2015). Bioengineering strategies on catalysis for the effective production of renewable and sustainable energy. *Renewable and Sustainable Energy Reviews*, *51*, 533–547. doi:10.1016/j.rser.2015.06.030

Harjunpää, V., Helin, J., Koivula, A., Siika-aho, M., & Drakenberg, T. (1999). A comparative study of two retaining enzymes of Trichoderma reesei: transglycosylation of oligosaccharides catalysed by the cellobiohydrolase I, Cel7A, and the beta-mannanase, Man5A. *FEBS Letters*, *443*(2), 149–153. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9989594

Harmsen, P., Huijgen, W., Bermudez Lopez, L. M., & Bakker, R. R. C. (2010). *Literature review of physical and chemical pretreatment processes for lignocellulosic biomass*. Wageningen.

Hartmann, A., Baldani, J., Kirchhof, G., Aßmus, B., Hutzler, P., Springer, N., … Döbereiner, J. (1995). Taxonomic and Ecologic Studies of Diazotrophic Rhizosphere Bacteria Using Phylogenetic Probes. In I. Fendrik, M. del Gallo, J. Vanderleyden, & M. de Zamaroczy (Eds.), *Azospirillum VI and Related Microorganisms SE - 47* (Vol. 37, pp. 415–427). Springer Berlin Heidelberg. doi:10.1007/978-3-642-79906-8_47

He, J., Yu, B., Zhang, K., Ding, X., & Chen, D. (2009). Expression of endo-1, 4-beta-xylanase from Trichoderma reesei in Pichia pastoris and functional characterization of the produced enzyme. *BMC Biotechnology*, *9*, 56. doi:10.1186/1472-6750-9-56

He, Y., Pang, Y., Liu, Y., Li, X., & Wang, K. (2008). Physicochemical Characterization of Rice Straw Pretreated with Sodium Hydroxide in the Solid State for Enhancing Biogas Production. *Energy & Fuels*, *22*(4), 2775–2781. doi:10.1021/ef8000967

Henco, K. (2003). *The QIAexpressionist*. *Diagen GmbH, Diisseldorf, Germany* (5th ed.). Redwood City: QIAGEN. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+QIA+expressionist+TM#0\nhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+QIA+expressionist#0

Hendriks, a T. W. M., & Zeeman, G. (2009). Pretreatments to enhance the digestibility of lignocellulosic biomass. *Bioresource Technology*, *100*(1), 10–8. doi:10.1016/j.biortech.2008.05.027

Herrera-González, V. E., Ruiz-Ordaz, N., Galíndez-Mayer, J., Juárez-Ramírez, C., Santoyo-Tepole, F., & Montiel, E. M. (2013). Biodegradation of the herbicide propanil, and its 3,4-dichloroaniline by-product in a continuously operated biofilm reactor. *World Journal of Microbiology & Biotechnology*, *29*(3), 467–74. doi:10.1007/s11274-012-1200-5

Hesse, S., Liebert, T., & Heinze, T. (2006). Studies on the film formation of

polysaccharide based furan-2-carboxylic acid esters. *Macromolecular Symposia*, *232*, 57–67. doi:10.1002/masy.200551407

Himmel, M. E., Xu, Q., Luo, Y., Ding, S.-Y., Lamed, R., & Bayer, E. a. (2010). Microbial enzyme systems for biomass conversion: emerging paradigms. *Biofuels*, *1*(2), 323–341. doi:10.4155/bfs.09.25

Hmidet, N., Bayoudh, A., Berrin, J. G., Kanoun, S., Juge, N., & Nasri, M. (2008). Purification and biochemical characterization of a novel α-amylase from Bacillus licheniformis NH1. *Process Biochemistry*, *43*(5), 499–510. doi:10.1016/j.procbio.2008.01.017

Hoben, H. J., & Somasegaran, P. (1982). Comparison of the Pour, Spread, and Drop Plate Methods for Enumeration of Rhizobium spp. in Inoculants Made from Presterilized Peat. *Applied and Environmental Microbiology*, *44*(5), 1246–1247.

Hodgson, E. M., Nowakowski, D. J., Shield, I., Riche, a, Bridgwater, a V, Clifton-Brown, J. C., & Donnison, I. S. (2011). Variation in Miscanthus chemical composition and implications for conversion by pyrolysis and thermo-chemical bio-refining for fuels and chemicals. *Bioresource Technology*, *102*(3), 3411–8. doi:10.1016/j.biortech.2010.10.017

Holt, J. G., Krieg, N. R., Sneath, P. H. A., Staley, J. T., & Williams, S. T. (2000). *Bergey's Manual of Determinative Bacteriology*. (W. R. Hensyl, Ed.) (9th ed.). Philidelphia: Lippincott Williams & Wilkins.

Holzman, T. F., Fesik, S. W., Park, C., & Kofron, J. L. (1991). Isolation and Characterization of Natural and Recombinant Cyclophilins. In J. Kelly & T. Baldwin (Eds.), *Applications of enzyme biotechnology* (1st ed., p. 115). New York: Plenum Press. doi:10.1007/978-1-4757-9235-5

Honda, Y., & Kitaoka, M. (2004). A family 8 glycoside hydrolase from Bacillus halodurans C-125 (BH2105) is a reducing end xylose-releasing exo-oligoxylanase. *The Journal of Biological Chemistry*, *279*(53), 55097–103. doi:10.1074/jbc.M409832200

Hoque, S. N., Graham, J., Kaufmann, M. E., & Tabaqchali, S. (2001). Chryseobacterium (Flavobacterium) meningosepticum outbreak associated with colonization of water taps in a neonatal intensive care unit. *Journal of Hospital Infection*, *47*(3), 188–192. doi:10.1053/jhin.2000.0908

Horikoshi, K. (1996). Alkaliphiles — from an industrial point of view. *FEMS Microbiology Reviews*, *18*(2-3), 259–270. doi:10.1111/j.1574-6976.1996.tb00242.x

Horn, S. J., Sørbotten, A., Synstad, B., Sikorski, P., Sørlie, M., Vårum, K. M., & Eijsink, V. G. H. (2006). Endo/exo mechanism and processivity of family 18 chitinases produced by Serratia marcescens. *The FEBS Journal*, *273*(3), 491–503. doi:10.1111/j.1742-4658.2005.05079.x

Hu, J., Arantes, V., & Saddler, J. N. (2011). The enhancement of enzymatic hydrolysis of lignocellulosic substrates by the addition of accessory enzymes such as xylanase: is it an additive or synergistic effect? *Biotechnology for Biofuels*, *4*(1), 36.

doi:10.1186/1754-6834-4-36

Huang, R., Su, R., Qi, W., & He, Z. (2011). Bioconversion of Lignocellulose into Bioethanol: Process Intensification and Mechanism Research. *BioEnergy Research*, *4*(4), 225–245. doi:10.1007/s12155-011-9125-7

Igbokwe, G. E., Ngobidi, K. C., & Iwuchukwu, N. P. (2013). Production of alpha-amylase from mixed Actinomycetes spp cultures at room temperature using Nelsons Colorimetric method. *Pakistan Journal of Biological Sciences*. doi:10.3923/pjbs.2013

Ingram, L., Gomez, P., Lai, X., Moniruzzaman, M., Wood, B., Yomano, L., & York, S. (1998). Metabolic engineering of bacteria for ethanol production. *Biotechnology and Bioengineering*, *58*(2-3), 204–14. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10191391

Invitrogen. (2010). *Champion pET Directional TOPO Experssion Kits User Manual*. Paisley.

Ito, K., Ito, S., Shimamura, T., Weyand, S., Kawarasaki, Y., Misaka, T., … Iwata, S. (2011). Crystal Structure of Glucansucrase from the Dental Caries Pathogen Streptococcus mutans. *Journal of Molecular Biology*, *408*(2), 177–186. doi:10.1016/j.jmb.2011.02.028

Jabbour, D., Borrusch, M. S., Banerjee, G., & Walton, J. D. (2013). Enhancement of fermentable sugar yields by α-xylosidase supplementation of commercial cellulases. *Biotechnology for Biofuels*, *6*(1), 58. doi:10.1186/1754-6834-6-58

Jaenicke, R. (1991). Protein stability and molecular adaptation to extreme conditions. *European Journal of Biochemistry / FEBS*, *202*(3), 715–728. doi:10.1111/j.1432-1033.1991.tb16426.x

Jaenicke, R., & Závodszky, P. (1990). Proteins under extreme physical conditions. *FEBS Letters*, *268*(2), 344–349. doi:10.1016/0014-5793(90)81283-T

Janda, J. M., & Abbott, S. L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *Journal of Clinical Microbiology*, *45*(9), 2761–2764. doi:10.1128/JCM.01228-07

Janssen, P. H., Yates, P. S., Grinton, B. E., Taylor, P. M., & Sait, M. (2002). Improved Culturability of Soil Bacteria and Isolation in Pure Culture of Novel Members of the Divisions Acidobacteria, Actinobacteria, Proteobacteria, and Verrucomicrobia. *Applied and Environmental Microbiology*, *68*(5), 2391–2396. doi:10.1128/AEM.68.5.2391

Jantama, K., Haupt, M. J., Svoronos, S. a., Zhang, X., Moore, J. C., Shanmugam, K. T., & Ingram, L. O. (2008). Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of Escherichia coli C that produce succinate and malate. *Biotechnology and Bioengineering*, *99*(5), 1140–1153. doi:10.1002/bit.21694

Jarboe, L. R., Zhang, X., Wang, X., Moore, J. C., Shanmugam, K. T., & Ingram, L. O. (2010). Metabolic engineering for production of biorenewable fuels and

chemicals: contributions of synthetic biology. *Journal of Biomedicine & Biotechnology*, *2010*, 761042. doi:10.1155/2010/761042

Javier, E., Barriga, C., Guamán-burneo, C., Portero, P., Salas, E., Tufiño, C., & Bastidas, B. (2013). Second Generation Ethanol from Residual Biomass : Research and Perspectives in Ecuador. In M. D. Matovic (Ed.), *Biomass Now - Sustainable Growth and Use* (1st ed., pp. 265–284). Rijeka: InTech. doi:10.5772/54528

Jeffries, T. W. (2006). Engineering yeasts for xylose metabolism. *Current Opinion In Biotechnology*, *17*(3), 320–6. doi:10.1016/j.copbio.2006.05.008

Jeng, W.-Y., Wang, N.-C., Lin, M.-H., Lin, C.-T., Liaw, Y.-C., Chang, W.-J., … Wang, A. H.-J. (2011). Structural and functional analysis of three β-glucosidases from bacterium Clostridium cellulovorans, fungus Trichoderma reesei and termite Neotermes koshunensis. *Journal of Structural Biology*, *173*(1), 46–56. doi:10.1016/j.jsb.2010.07.008

Jeon, B.-H., Choi, J.-A., Kim, H.-C., Hwang, J.-H., Abou-Shanab, R. A., Dempsey, B. a, … Kim, J. R. (2013). Ultrasonic disintegration of microalgal biomass and consequent improvement of bioaccessibility/bioavailability in microbial fermentation. *Biotechnology for Biofuels*, *6*(1), 37. doi:10.1186/1754-6834-6-37

Jiang, J., Alderisio, K. A., Singh, A., Xiao, L., & Icrobiol, A. P. P. L. E. N. M. (2005). Development of Procedures for Direct Extraction of Cryptosporidium DNA from Water Concentrates and for Relief of PCR Inhibitors. *Applied and Environmental Microbiology*, *71*(3), 1135–1141. doi:10.1128/AEM.71.3.1135

Johansen, J. E., Binnerup, S. J., Kroer, N., & Mølbak, L. (2005). Luteibacter rhizovicinus gen. nov., sp. nov., a yellow-pigmented gammaproteobacterium isolated from the rhizosphere of barley (Hordeum vulgare L.). *International Journal of Systematic and Evolutionary Microbiology*, *55*(Pt 6), 2285–91. doi:10.1099/ijs.0.63497-0

Johnson, B., Markham, T., Samyoylov, V., & Dallmier, K. (2006). Corn event 3272 and methods of detection thereof. U.S.A: World Intellectual Property Organization International Beureau. Retrieved from www.google.com/patents/WO2006098952

Jones, R. M., Hedrich, S., & Johnson, D. B. (2013). Acidocella aromatica sp. nov.: An acidophilic heterotrophic alphaproteobacterium with unusual phenotypic traits. *Extremophiles*, *17*(5), 841–850. doi:10.1007/s00792-013-0566-0

Jüergensen, J., Ilmberger, N., & Streit, W. R. (2012). Microbial Metabolic Engineering: Methods and Protocols. In Q. Cheng (Ed.), *Methods in Molecular Biology* (834th ed., Vol. 834, pp. 1–16). New York, NY: Springer New York. doi:10.1007/978-1-61779-483-4

Jung, H.-M., Ten, L. N., Kim, K.-H., An, D. S., Im, W.-T., & Lee, S.-T. (2009). Dyella ginsengisoli sp. nov., isolated from soil of a ginseng field in South Korea. *International Journal of Systematic and Evolutionary Microbiology*, *59*(Pt 3), 460–5. doi:10.1099/ijs.0.64514-0

Juturu, V., & Wu, J. C. (2014). Microbial Exo-xylanases: A Mini Review. *Applied Biochemistry and Biotechnology*, *174*(1), 81–92. doi:10.1007/s12010-014-1042-8

Kamble, R. D., & Jadhav, A. R. (2012). Isolation, purification, and characterization of xylanase produced by a new species of bacillus in solid state fermentation. *International Journal of Microbiology*, *2012*. doi:10.1155/2012/683193

Kamm, B., & Kamm, M. (2004a). Biorefinery – Systems. *Chemical and Biochemical Engineering Quarterly*, *18*(1), 1–6.

Kamm, B., & Kamm, M. (2004b). Principles of biorefineries. *Applied Microbiology and Biotechnology*, *64*(2), 137–45. doi:10.1007/s00253-003-1537-7

Kampmann, M., & Stock, D. (2004). Reverse gyrase has heat-protective DNA chaperone activity independent of supercoiling. *Nucleic Acids Research*, *32*(12), 3537–3545. doi:10.1093/nar/gkh683

Kanehisa Laboratories. (2015). KEGG Automatic Annotation Server.

Kasana, R. C., Salwan, R., Dhar, H., Dutt, S., & Gulati, A. (2008). A rapid and easy method for the detection of microbial cellulases on agar plates using gram's iodine. *Current Microbiology*, *57*(5), 503–507. doi:10.1007/s00284-008-9276-8

Kataeva, I. a., Yang, S. J., Dam, P., Poole, F. L., Yin, Y., Zhou, F., … Adams, M. W. W. (2009). Genome sequence of the anaerobic, thermophilic, and cellulolytic bacterium "Anaerocellum thermophilum" DSM 6725. *Journal of Bacteriology*, *191*(11), 3760–3761. doi:10.1128/JB.00256-09

Kazenwadel, F., Franzreb, M., & Rapp, B. E. (2015). Synthetic Enzyme Supercomplexes: Co-Immobilization of Enzyme Cascades. *Anal. Methods*, *7*, 4030–4037. doi:10.1039/C5AY00453E

Kelley, L. a, Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, *10*(6), 845–858. doi:10.1038/nprot.2015-053

Kelley, L. a, & Sternberg, M. J. E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols*, *4*(3), 363–71. doi:10.1038/nprot.2009.2

Kennedy, J., Margassery, L. M., Morrissey, J. P., O'Gara, F., & Dobson, A. D. W. (2013). Sequenced-based screening: metagenomic DNA. In A. Trincone (Ed.), *Marine Enzymes for Biocatalysis: Sources, Biocatalytic Characteristics and Bioprocesses of Marine Enzymes* (1st ed., pp. 115–117). Cambridge: Woodhead Publishing Limited.

Kenyon College. (2010). Bacillus. Retrieved from http://microbewiki.kenyon.edu/index.php/Bacillus

Khasin, A., Alchanati, I., & Shoham, Y. (1993). Purification and characterization of a thermostable xylanase from Bacillus stearothermophilus T-6. *Applied and Environmental Microbiology*, *59*(6), 1725–1730.

Khurshed, N. (2003). *Studies on the locally isolated strains of Bacillus for their possible use as bioinsecticides*. University of the Punjab. Retrieved from http://prr.hec.gov.pk/Thesis/1394.pdf

Kim, D. Y., Ham, S.-J., Lee, H. J., Cho, H.-Y., Kim, J.-H., Kim, Y.-J., … Park, H.-Y. (2011). Cloning and characterization of a modular GH5 β-1,4-mannanase with high specific activity from the fibrolytic bacterium Cellulosimicrobium sp. strain HY-13. *Bioresource Technology*, *102*(19), 9185–92. doi:10.1016/j.biortech.2011.06.073

Kim, T. H., & Lee, Y. Y. (2005). Pretreatment of Corn Stover by Soaking in aqueous ammonia. *Applied Biochemistry and Biotechnology*, *121-124*, 1119–1131.

Kim, Y. S., Jung, H. C., & Pan, J. G. (2000). Bacterial cell surface display of an enzyme library for selective screening of improved cellulase variants. *Applied and Environmental Microbiology*, *66*(2), 788–93. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=91897&tool=pmcentrez&rendertype=abstract

Kimura, Z., Chung, K. M., Itoh, H., Hiraishi, A., & Okabe, S. (2014). Raoultella electrica sp. nov., isolated from anodic biofilms of a glucose-fed microbial fuel cell. *International Journal of Systematic and Evolutionary Microbiology*, *64*(Pt 4), 1384–1388. doi:10.1099/ijs.0.058826-0

Kind, T., Meissen, J. K., Yang, D., Nocito, F., Vaniya, A., Cheng, Y.-S., … Fiehn, O. (2012). Qualitative analysis of algal secretions with multiple mass spectrometric platforms. *Journal of Chromatography A*, *1244*, 139–147. doi:10.1016/j.biotechadv.2011.08.021.Secreted

King, B. C., Donnelly, M. K., Bergstrom, G. C., Walker, L. P., & Gibson, D. M. (2009). An optimized microplate assay system for quantitative evaluation of plant cell wall-degrading enzyme activity of fungal culture extracts. *Biotechnology and Bioengineering*, *102*(4), 1033–44. doi:10.1002/bit.22151

Kirchhof, G., Eckert, B., Stoffels, M., Baldani, J. I., Reis, V. M., & Hartmann, a. (2001). Herbaspirillum frisingense sp. nov., a new nitrogen-fixing bacterial species that occurs in C4-fibre plants. *International Journal of Systematic and Evolutionary Microbiology*, *51*(Pt 1), 157–68. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11211253

Kirchhof, G., Reis, V. M., Baldani, J. I., Eckert, B., & Hartmann, A. (1997). Occurrence, physiological and molecular analysis of endophytic diazotrophic bacteria in gramineous energy plants. *Plant and Soil*, *194*, 45–55.

Klamer, M., & Bååth, E. (1998). Microbial community dynamics during composting of straw material studied using phospholipid fatty acid analysis. *FEMS Microbiology Ecology*, *27*, 9–20. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1574-6941.1998.tb00521.x/full

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., & Glöckner, F. O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research*, *41*(1), 1–11. doi:10.1093/nar/gks808

Kloepper, J. W., Ryu, C.-M., & Zhang, S. (2004). Induced Systemic Resistance and Promotion of Plant Growth by Bacillus spp. *Phytopathology*, *94*(11), 1259–66. doi:10.1094/PHYTO.2004.94.11.1259

Ko, C.-H., Chen, W.-L., Tsai, C.-H., Jane, W.-N., Liu, C.-C., & Tu, J. (2007). Paenibacillus campinasensis BL11: a wood material-utilizing bacterial strain isolated from black liquor. *Bioresource Technology*, *98*(14), 2727–33. doi:10.1016/j.biortech.2006.09.034

Koeck, D. E., Pechtl, A., Zverlov, V. V., & Schwarz, W. H. (2014). Genomics of cellulolytic bacteria. *Current Opinion in Biotechnology*, *29*(1), 171–183. doi:10.1016/j.copbio.2014.07.002

Kohring, S., Wiegel, J., & Mayer, F. (1990). Subunit Composition and Glycosidic Activities of the Cellulase Complex from Clostridium thermocellum JW20. *Applied and Environmental Microbiology*, *56*(12), 3798–804. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=185070&tool=pmcentrez&rendertype=abstract

Kongjan, P., O-Thong, S., Kotay, M., Min, B., & Angelidaki, I. (2010). Biohydrogen production from wheat straw hydrolysate by dark fermentation using extreme thermophilic mixed culture. *Biotechnology and Bioengineering*, *105*(5), 899–908. doi:10.1002/bit.22616

Kongpol, A., Pongtharangkul, T., Kato, J., Honda, K., Ohtake, H., & Vangnai, A. S. (2009). Characterization of an organic-solvent-tolerant Brevibacillus agri strain 13 able to stabilize solvent/water emulsion. *FEMS Microbiology Letters*, *297*(2), 225–33. doi:10.1111/j.1574-6968.2009.01684.x

Koning, S. M., Albers, S.-V., Konings, W. N., & Driessen, A. J. M. (2002). Sugar transport in (hyper) thermophilic archaea. *Research in Microbiology*, *153*(2), 61–67. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11902154

Kordel, M., Hofmann, B., Schomburg, D., & Schmid, R. D. (1991). Extracellular lipase of Pseudomonas sp. strain ATCC 21808: Purification, characterization, crystallization, and preliminary X-ray diffraction data. *Journal of Bacteriology*, *173*(15), 4836–4841. doi:DS

Koyama, Y., Hoshino, T., Tomizuka, N., & Furukawa, K. (1986). Genetic Transformation of the Extreme Thermophile Thermus thermophilus and of Other Thermus spp. *Journal of Bacteriology*, *166*(1), 338–340.

Koyanagi, K. O., Hagiwara, M., Itoh, T., Gojobori, T., & Imanishi, T. (2005). Comparative genomics of bidirectional gene pairs and its implications for the evolution of a transcriptional regulation system. *Gene*, *353*(2), 169–176. doi:10.1016/j.gene.2005.04.027

Kulichevskaya, I. S., Detkova, E. N., Bodelier, P. L. E., Rijpstra, W. I. C., Damste, J. S. S., & Dedysh, S. N. (2011). Singulisphaera rosea sp. nov., a planctomycete from acidic sphagnum peat, and emended description of the genus singulisphaera. *International Journal of Systematic and Evolutionary Microbiology*, *62*(1), 118–123. doi:10.1099/ijs.0.025924-0

Kulichevskaya, I. S., Ivanova, A. O., Baulina, O. I., Bodelier, P. L. E., Damsté, J. S. S., & Dedysh, S. N. (2008). Singulisphaera acidiphila gen. nov., sp. nov., a non-filamentous, Isosphaera-like planctomycete from acidic northen wetlands.

International Journal of Systematic and Evolutionary Microbiology, *58*(5), 1186–1193. doi:10.1099/ijs.0.65593-0

Kumar, P., Barrett, D. M., Delwiche, M. J., & Stroeve, P. (2009). Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production. *Industrial & Engineering Chemistry Research*, *48*(8), 3713–3729. doi:10.1021/ie801542g

Kunst, F., Ogasawara, N., Moszer, I., Albertini, a M., Alloni, G., Azevedo, V., … Danchin, a. (1997). The complete genome sequence of the gram-positive bacterium Bacillus subtilis. *Nature*, *390*(6657), 249–256. doi:10.1038/36786

La Grange, D. C., Pretorius, I. S., & Claeyssens, M. (2001). Degradation of xylan to D -xylose by recombinant Saccharomyces cerevisiae coexpressing the Aspergillus niger b -xylosidase (xlnD) and the Trichoderma reesei xylanase II (xyn2) genes. *Applied and Environmental Microbiology*, *67*(12), 5512–5519. doi:10.1128/AEM.67.12.5512

Lamosa, P., Burke, A., Peist, R., Huber, R., Liu, M. Y., Silva, G., … Santos, H. (2000). Thermostabilization of proteins by diglycerol phosphate, a new compatible solute from the hyperthermophile Archaeoglobus fulgidus. *Applied and Environmental Microbiology*, *66*(5), 1974–1979. doi:10.1128/AEM.66.5.1974-1979.2000

Langeveld, J. W. A., de Jong, E., van Ree, R., & Sanders, J. P. M. (2010). Biorefinery: Definition and classification system. In J. W. A. Langeveld, M. Meeusen, & J. Sanders (Eds.), *The biobased economy biofuels, materials and chemicals in the post-oil era* (1st ed., p. 114). London: Earthscan. Retrieved from http://books.google.co.uk/books?id=I6YIT1ElwLoC&printsec=frontcover&dq=The+biobased+economy:+biofuels,+materials+and+chemicals+in+the+post-oil+era&hl=en&sa=X&ei=CY-sUcHiNIWyOb2LgcgD&ved=0CDQQ6AEwAA#v=onepage&q=The biobased economy: biofuels, materials and chemicals in the post-oil era&f=false

Larsson, S., Palmqvist, E., Hahn-Hägerdal, B., Tengborg, C., Stenberg, K., Zacchi, G., & Nilvebrant, N.-O. (1999). The generation of fermentation inhibitors during dilute acid hydrolysis of softwood. *Enzyme and Microbial Technology*, *24*(3-4), 151–159. doi:10.1016/S0141-0229(98)00101-X

Lavarack, B. P., Griffin, G. J., & Rodman, D. (2002). The acid hydrolysis of sugarcane bagasse hemicellulose to produce xylose, arabinose, glucose and other products. *Biomass and Bioenergy*, *23*, 367–380.

Le, N. T., Julcour-Lebigue, C., & Delmas, H. (2013). Ultrasonic sludge pretreatment under pressure. *Ultrasonics Sonochemistry*, *20*(5), 1203–10. doi:10.1016/j.ultsonch.2013.03.005

Leandro, M. J., Gonçalves, P., & Spencer-Martins, I. (2006). Two glucose/xylose transporter genes from the yeast Candida intermedia: first molecular characterization of a yeast xylose-H+ symporter. *The Biochemical Journal*, *395*(3), 543–9. doi:10.1042/BJ20051465

Lee, C. C., Wong, D. W., & Robertson, G. H. (2001). An E. coli expression system for the

extracellular secretion of barley alpha-amylase. *Journal of Protein Chemistry*, *20*(3), 233–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11565903

Lee, C. Y., Bagdasarian, M., & Zeikus, J. G. (1990). Catalytic mechanism of xylose (glucose) isomerase from Clostridium thermosulfurogenes. Characterization of the structural gene and function of active site histidine. *The Journal of Biological Chemistry*, *265*, 19082–19090.

Lee, C., & Zeikus, J. G. (1991). Purification and characterization of thermostable glucose isomerase from Clostridium thermosulfurogenes and Thermoanaerobacter strain B6A. *Biochemical Journal*, *273*, 565–571.

Lee, D., Owens, V. N., Boe, A., & Jeranyama, P. (2007). *Composition of Herbaceous Biomass Feedstocks*. Brookings.

Lee, J. C., & Gutell, R. R. (2012). A comparison of the crystal structures of eukaryotic and bacterial SSU ribosomal RNAs reveals common structural features in the hypervariable regions. *PloS One*, *7*(5), e38203. doi:10.1371/journal.pone.0038203

Lee, J., Jang, Y. S., Choi, S. J., Im, J. A., Song, H., Cho, J. H., … Lee, S. Y. (2012). Metabolic engineering of clostridium acetobutylicum ATCC 824 for isopropanol-butanol-ethanol fermentation. *Applied and Environmental Microbiology*, *78*(5), 1416–1423. doi:10.1128/AEM.06382-11

Lee, J.-W., Kim, H.-Y., Koo, B.-W., Choi, D.-H., Kwon, M., & Choi, I.-G. (2008). Enzymatic saccharification of biologically pretreated Pinus densiflora using enzymes from brown rot fungi. *Journal of Bioscience and Bioengineering*, *106*(2), 162–7. doi:10.1263/jbb.106.162

Lee, S. E., Armiger, W. B., Watteeuw, C. M., & Humphrey, A. E. (1978). A Theoretical Model for Enzymatic Hydrolysis o f Cellulose. *Biotechnology and Bioengineering*, *20*(1), 141–144.

Lee, S. F., Forsberg, C. W., & Gibbins, L. N. (1985). Xylanolytic activity of Clostridium acetobutylicum. *Applied Environmental Microbiology*, *50*(4), 1068–1076.

Lee, S. J., Warnick, T. a, Pattathil, S., Alvelo-Maurosa, J. G., Serapiglia, M. J., McCormick, H., … Hazen, S. P. (2012). Biological conversion assay using Clostridium phytofermentans to estimate plant feedstock quality. *Biotechnology for Biofuels*, *5*(1), 5. doi:10.1186/1754-6834-5-5

Lee, S. J., Warnick, T. A., Leschine, S. B., & Hazen, S. P. (2012). High-throughput biological conversion assay for determining lignocellulosic quality. In J. Normanly (Ed.), *High-Throughput Phenotyping in Plants: Methods and Protocols* (1st ed., Vol. 918, pp. 341–349). Totowa, NJ: Humana Press. doi:10.1007/978-1-61779-995-2

Lee, S. J., Warnick, T. A., Pattathil, S., Alvelo-maurosa, J. G., Serapiglia, M. J., Mccormick, H., … Hazen, S. P. (2012). Biological conversion assay using Clostridium phytofermentans to estimate plant feedstock quality. *Biotechnology for Biofuels*, *5*(February), 5–18.

Lee, S.-H., Kim, S.-J., & Yi, B.-Y. (2009). Simplex and duplex event-specific analytical

methods for functional biotech maize. *Journal of Agricultural and Food Chemistry*, *57*(16), 7178–7185. doi:10.1021/jf901078d

Lewin, B. (2007). The nucleus. In B. Lewin, L. Cassimeris, V. R. Lingappa, & G. Plopper (Eds.), *Cells* (First., p. 863). London: Jones & Bartlett Learning.

Li, D. (2011). *ECOLOGY OF NITROGEN-FIXING BACTERIA ASSOCIATED WITH MISCANTHUS*. Graduate College of the University of Illinois.

Li, J., Dai, J., Chen, X., & Zhu, P. (2007). Microbial Transformation of Cephalomannine by Luteibacter sp . *Journal of Natural Products*, *70*, 1846–1849. doi:10.1021/np0701531

Liang, Y., Yesuf, J., Schmitt, S., Bender, K., & Bozzola, J. (2009). Study of cellulases from a newly isolated thermophilic and cellulolytic Brevibacillus sp. strain JXL. *Journal of Industrial Microbiology & Biotechnology*, *36*(7), 961–70. doi:10.1007/s10295-009-0575-2

Linger, J. G., Adney, W. S., & Darzins, A. (2010). Heterologous expression and extracellular secretion of cellulolytic enzymes by Zymomonas mobilis. *Applied and Environmental Microbiology*, *76*(19), 6360–9. doi:10.1128/AEM.00230-10

Linke, C. M., Woodiga, S. a., Meyers, D. J., Buckwalter, C. M., Salhi, H. E., & King, S. J. (2013). The ABC transporter encoded at the pneumococcal fructooligosaccharide utilization locus determines the ability to utilize long- and short-chain fructooligosaccharides. *Journal of Bacteriology*, *195*(5), 1031–1041. doi:10.1128/JB.01560-12

Lionetti, V., Francocci, F., Ferrari, S., Volpi, C., Bellincampi, D., Galletti, R., … Cervone, F. (2010). Engineering the cell wall by reducing de-methyl-esterified homogalacturonan improves saccharification of plant tissues for bioconversion. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(2), 616–21. doi:10.1073/pnas.0907549107

Liu, C. H., Lu, W. Bin, & Chang, J. S. (2006). Optimizing lipase production of Burkholderia sp. by response surface methodology. *Process Biochemistry*, *41*(9), 1940–1944. doi:10.1016/j.procbio.2006.04.013

Liu, J., Sun, D., Liu, H., Nie, Y., & Zhu, Z. (2010). Biodegradation Kinetics for Pre-treatment of Klebsiella pneumoniae Waste with Autothermal Thermophilic Aerobic Digestion. *Chinese Journal of Chemical Engineering*, *18*(6), 905–909. doi:10.1016/S1004-9541(09)60146-4

Liu, L., Liu, Y., Shin, H. D., Chen, R. R., Wang, N. S., Li, J., … Chen, J. (2013). Developing Bacillus spp. as a cell factory for production of microbial enzymes and industrially important biochemicals in the context of systems and synthetic biology. *Applied Microbiology and Biotechnology*, *97*(14), 6113–6127. doi:10.1007/s00253-013-4960-4

Liu, N., Yan, X., Zhang, M., Xie, L., Wang, Q., Huang, Y., … Zhou, Z. (2011). Microbiome of Fungus-Growing Termites : a New Reservoir for Lignocellulase Genes. *Applied and Environmental Microbiology*, *77*(1), 48–56. doi:10.1128/AEM.01521-10

Liu, Y., Zhang, J., Liu, Q., Zhang, C., & Ma, Q. (2004). Molecular cloning of novel cellulase genes cel9A and cel12A from Bacillus licheniformis GXN151 and synergism of their encoded polypeptides. *Current Microbiology*, *49*(4), 234–8. doi:10.1007/s00284-004-4291-x

Livesey, G. (2003). Health potential of polyols as sugar replacers, with emphasis on low glycaemic properties. *Nutrition Research Reviews*, *16*(2), 163–91. doi:10.1079/NRR200371

Lloberas, J., Querol, E., & Bernues, J. (1988). Purification and characterisation of endo-b-13-14-D-glucanase activity from Bacillus licheniformis. *Applied Microbiology and Biotechnology*, *29*, 32–38.

Löffler, F. E., Sun, Q., Li, J., & Tiedje, J. M. (2000). 16S rRNA gene-based detection of tetrachloroethene-dechlorinating Desulfuromonas and Dehalococcoides species. *Applied and Environmental Microbiology*, *66*(4), 1369–1374. doi:10.1128/AEM.66.4.1369-1374.2000

Logan, N. a, Forsyth, G., Lebbe, L., Goris, J., Heyndrickx, M., Balcaen, a, … De Vos, P. (2002). Polyphasic identification of Bacillus and Brevibacillus strains from clinical, dairy and industrial specimens and proposal of Brevibacillus invocatus sp. nov.. *International Journal of Systematic and Evolutionary Microbiology*, *52*(Pt 3), 953–66. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12054263

Lönn, A., Träff-Bjerre, K. ., Cordero Otero, R. ., van Zyl, W. ., & Hahn-Hägerdal, B. (2003). Xylose isomerase activity influences xylose fermentation with recombinant Saccharomyces cerevisiae strains expressing mutated xylA from Thermus thermophilus. *Enzyme and Microbial Technology*, *32*(5), 567–573. doi:10.1016/S0141-0229(03)00024-3

Lowe, S. E., Jain, M. K., & Zeikus, J. G. (1993). Biology, ecology, and biotechnological applications of anaerobic bacteria adapted to environmental stresses in temperature, pH, salinity, or substrates. *Microbiological Reviews*, *57*(2), 451–509. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=372919&tool=pmcentrez&rendertype=abstract

Lütke-Eversloh, T., & Bahl, H. (2011). Metabolic engineering of Clostridium acetobutylicum: Recent advances to improve butanol production. *Current Opinion in Biotechnology*, *22*(5), 634–647. doi:10.1016/j.copbio.2011.01.011

Lygin, A. V., Upton, J., Dohleman, F. G., Juvik, J., Zabotina, O. a., Widholm, J. M., & Lozovaya, V. V. (2011). Composition of cell wall phenolics and polysaccharides of the potential bioenergy crop -Miscanthus. *GCB Bioenergy*, *3*(4), 333–345. doi:10.1111/j.1757-1707.2011.01091.x

Lymar, E. S., Li, B., & Renganathan, V. (1995). Purification and Characterization of a Cellulose-Binding (beta)-Glucosidase from Cellulose-Degrading Cultures of Phanerochaete chrysosporium. *Applied and Environmental Microbiology*, *61*(8), 2976–2980.

Lynd, L. R., Weimer, P. J., Zyl, W. H. Van, & Pretorius, I. S. (2002). Microbial Cellulose

Utilization: Fundamentals and Biotechnology. *Microbiology and Molecular Biology Reviews*, *66*(3), 506–577. doi:10.1128/MMBR.66.3.506

Madigan, M. T., & Oren, A. (1999). Thermophilic and halophilic extremophiles. *Current Opinion in Microbiology*, *2*(3), 265–269. doi:10.1016/S1369-5274(99)80046-0

Maki, M., Leung, K. T., & Qin, W. (2009). The prospects of cellulase-producing bacteria for the bioconversion of lignocellulosic biomass. *International Journal of Biological Sciences*, *5*(5), 500–16. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2726447&tool=pmcentrez&rendertype=abstract

Malhotra, R., Noorwez, S. M., & Satyanarayana, T. (2000). Production and partial characterization of thermostable and calcium-independent a-amylase of an extreme thermophile Bacillus thermooleovorans NP54. *Letters in Applied Microbiology*, *31*, 378–384.

Mansfield, S., Mooney, C., & Saddler, J. (1999). Substrate and Enzyme Characteristics that Limit Cellulose Hydrolysis. *Biotechnology Progress*, *15*(5), 804–816. doi:10.1021/bp9900864

Map of Life. (2012). Extremophiles Archaea and Bacteria. Retrieved March 26, 2015, from http://www.mapoflife.org/topics/topic_354_Extremophiles-Archaea-and-Bacteria/

Marchesi, J. R., Sato, T., Weightman, A. J., Martin, T. a., Fry, J. C., Hiom, S. J., & Wade, W. G. (1998). Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. *Applied and Environmental Microbiology*, *64*(2), 795–799.

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., … Bryant, S. H. (2014). CDD: NCBI's conserved domain database. *Nucleic Acids Research*, *43*(D1), D222–D226. doi:10.1093/nar/gku1221

Martinez, D., Berka, R. M., Henrissat, B., Saloheimo, M., Arvas, M., Baker, S. E., … Brettin, T. S. (2008). Genome sequencing and analysis of the biomass-degrading fungus Trichoderma reesei (syn. Hypocrea jecorina). *Nature Biotechnology*, *26*(5), 553–560. doi:10.1038/nbt1403

Martins, D., Prado, H. Do, Leite, R., Ferreira, H., Moretti, M. M. de S., da Silva, R., & Gomes, E. (2011). Agroindustrial wastes as substrates for microbial enzymes production and source of sugar for bioethanol production. In S. Kumar (Ed.), *Integrated Waste Management Volume II* (p. 482). InTech. doi:10.5772/23377

Martoglio, B., & Dobberstein, B. (1998). Signal sequences: more than just greasy peptides. *Trends in Cell Biology*, *8*(10), 410–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9789330

Mata-Alvarez, J., Macé, S., & Llabrés, P. (2000). Anaerobic digestion of organic solid wastes. An overview of research achievements ans perspectives. *Bioresource Technology*, *74*, 3–16.

Matthews, J. F., Bergenstråhle, M., Beckham, G. T., Himmel, M. E., Nimlos, M. R.,

Brady, J. W., & Crowley, M. F. (2011). High-temperature behavior of cellulose I. *The Journal of Physical Chemistry. B*, *115*(10), 2155–66. doi:10.1021/jp1106839

McCarthy, A., Peace, E., & Broda, P. (1985). Studies on the extracellular xylanase activity of some thermophilic actinomycetes. *Applied Microbiology and Biotechnology*, *21*(3-4), 238–244. doi:10.1007/BF00295129

McNeil, M. M., & Brown, J. M. (1994). The medically important aerobic actinomycetes: Epidemiology and microbiology. *Clinical Microbiology Reviews*, *7*(3), 357–417. doi:10.1128/CMR.7.3.357.Updated

Medema, M. H., Breitling, R., Bovenberg, R., & Takano, E. (2011). Exploiting plug-and-play synthetic biology for drug discovery and production in microorganisms. *Nature Reviews. Microbiology*, *9*(2), 131–137. doi:10.1038/nrmicro2478

Mehta, D., & Satyanarayana, T. (2013). Dimerization mediates thermo-adaptation, substrate affinity and transglycosylation in a highly thermostable maltogenic amylase of Geobacillus thermoleovorans. *PloS One*, *8*(9), e73612. doi:10.1371/journal.pone.0073612

Melasniewmi, H. (1987). Characterization of a-amylase and pullulanase activities of Clostridium thermohydrosulfuricum. *Biochemical Journal*, *197*, 193–197.

Melford. (2011). Yeast extract. Ipswich: Melford Laboratories Ltd.

Mergaert, J., Cnockaert, M. C., & Swings, J. (2002). Fulvimonas soli gen. nov., sp. nov., a y-proteobacterium isolated from soil after enrichment on acetylated starch plastic. *International Journal of Systematic and Evolutionary Microbiology*, *52*(4), 1285–1289. doi:10.1099/ijs.0.01995-0

Merritt, A. D., & Karn, R. C. (1977). The Human a-Amylases. In H. Harris & K. Hirschhorn (Eds.), *Advances in Human Genetics* (1st ed., pp. 159–160). New York: Plenum Press. doi:10.1007/978-1-4615-8269-4

Michels, M., & Bakker, E. P. (1985). Generation of a large, protonophore-sensitive proton motive force and pH difference in the acidophilic bacteria Thermoplasma acidophilum and Bacillus acidocaldarius. *Journal of Bacteriology*, *161*(1), 231–237.

Mielenz, J. (2009). *Biofuels: Methods and Protocols*. (J. R. Mielenz & J. M. Walker, Eds.) (1st ed.). New York: Humana Press.

Mishra, S., & Behera, N. (2008). Amylase activity of a starch degrading bacteria isolated from soil receiving kitchen wastes. *African Journal of Biotechnology*, *7*(18), 3326–3331.

Miyamoto, T., Kawahara, M., & Minamisawa, K. (2004). Novel Endophytic Nitrogen-Fixing Clostridia from the Grass Miscanthus sinensis as Revealed by Terminal Restriction Fragment Length Polymorphism Analysis. *Applied and Environmental Microbiology*, *70*(11), 6580–6586. doi:10.1128/AEM.70.11.6580

Mladenovska, Z., Mathrani, I. M., & Ahring, B. K. (1995). Isolation and characterization of Caldicellulosiruptor lactoaceticus sp. nov., an extremely thermophilic, cellulolytic, anaerobic bacterium. *Archives of Microbiology*, *163*(3), 223–230. doi:10.1007/BF00305357

Mohana, S., Shah, A., Divecha, J., & Madamwar, D. (2008). Xylanase production by Burkholderia sp. DMAX strain under solid state fermentation using distillery spent wash. *Bioresource Technology*, *99*(16), 7553–7564. doi:10.1016/j.biortech.2008.02.009

Mohanta, S., Sharma, G. D., & Deb, B. (2010). Diversity of Endophytic Diazotrophs in Non-Leguminous Crops - A Review. *Assam University Journal of Science & Technology: Biological and Environmental Sciences*, *6*(1), 109–122.

Morales, P., Madarro, a, Pérez-González, J. a, Sendra, J. M., Piñaga, F., & Flors, a. (1993). Purification and Characterization of Alkaline Xylanases from Bacillus polymyxa. *Applied and Environmental Microbiology*, *59*(5), 1376–82. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=182092&tool=pmcentrez&rendertype=abstract

MORRISON, W. R., & KARKALAS, J. (1990). Starch. In P. M. Dey & J. B. Harborne (Eds.), *Methods in plant biochemistry, Volume 2: carbohydrates* (1st ed., Vol. Volume 2, pp. 323–352). London: Academic Press. doi:http://dx.doi.org/10.1016/B978-0-12-461012-5.50015-X

Mosier, N., Hendrickson, R., Ho, N., Sedlak, M., & Ladisch, M. R. (2005). Optimization of pH controlled liquid hot water pretreatment of corn stover. *Bioresource Technology*, *96*(18), 1986–93. doi:10.1016/j.biortech.2005.01.013

Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y. Y., Holtzapple, M., & Ladisch, M. (2005). Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology*, *96*(6), 673–86. doi:10.1016/j.biortech.2004.06.025

Mukherjee, A. K., Borah, M., & Rai, S. K. (2009). To study the influence of different components of fermentable substrates on induction of extracellular α-amylase synthesis by Bacillus subtilis DM-03 in solid-state fermentation and exploration of feasibility for inclusion of α-amylase in laundry detergen. *Biochemical Engineering Journal*, *43*(2), 149–156. doi:10.1016/j.bej.2008.09.011

Mutondo, M. S., Huddy, R. J., Bauer, R., Tuffin, I. M., & Cowan, D. A. (2012). Meatgenomic gene dicovery. In R. W. Li (Ed.), *Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies* (1st ed., pp. 287–320). Hauppauge: Nova Science Publishers, Inc.

Nakamura, S., Nakai, R., Wakabayashi, K., Ishiguro, Y., Aono, R., & Horikoshi, K. (1994). Thermophilic alkaline xylanase from newly isolated alkaliphilic and thermophilic Bacillus sp. strain TAR-1. *Biosciences, Biotechnology and Biochemistry*, *58*(1), 78–81.

Nakayama, S., Kiyoshi, K., Kadokura, T., & Nakazato, A. (2011). Butanol production from crystalline cellulose by Cocultured Clostridium thermocellum and Clostridium saccharoperbutylacetonicum N1-4. *Applied and Environmental Microbiology*, *77*(18), 6470–6475. doi:10.1128/AEM.00706-11

National Center for Biotechnology Information. (2015). Genome Assembly and

Annotation report: Bacillus licheniformis. Retrieved July 17, 2015, from http://www.ncbi.nlm.nih.gov/genome/genomes/412?

National Microbial Pathogen Data Resource. (2015). Seed Viewer - Subsystem: Xylose utilisation. Retrieved April 13, 2015, from http://rast.nmpdr.org/seedviewer.cgi

NCBI. (2014). Bacillus coagulans Assemblies. Retrieved from http://www.ncbi.nlm.nih.gov/assembly/?term=bacillus coagulans

Nelson, D. L., & Cox, M. M. (2005). *Lehninger Principles of Biochemistry*. (L. Miller, Ed.) (4th ed.). New York: WH Freeman.

Neogen Corporation. (2011). YEAST EXTRACT (7184). Lansing, MI: Acumedia.

Niehaus, F., Bertoldo, C., Kähler, M., & Antranikian, G. (1999). Extremophiles as a source of novel enzymes for industrial application. *Applied Microbiology and Biotechnology*, *51*(6), 711–29. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10422220

Noller, H. F., Stolk, B. J. V. A. N., Douthwaite, S., & Gutell, R. R. (1985). Studies on the structure and function of 16S ribosomal RNA using structure-specific chemical probes. *Proc. Int. Symp. Biomol. Struct. Interactions, Suppl. J. Biosci*, *8*(August), 747–755.

Novel Foods Section. (2013). *Novel Food Information Alpha-amylase Corn Event 3272*. Ottawa. Retrieved from www.hc-sc.gc.ca/fn-an/gmf-agm/appro/nf-an144decdoc-eng.php

Novozymes. (2012, October). Novozymes and Terranol to market advanced biofuel yeast. doi:10.1016/S1351-4180(12)70396-3

Novozymes. (2013a). *Enzymes at work*. Bagsvaerd.

Novozymes. (2013b). World's first advanced biofuels facility opens. Retrieved from http://www.novozymes.com/en/news/news-archive/Pages/World's-first-advanced-biofuels-facility-opens.aspx

Novozymes Bioenergy. (2010a). *Cellic CTec2 and HTec2 - enzymes for hydrolysis of lignocellulosic materials*. Bagsvaerd.

Novozymes Bioenergy. (2010b). Novozymes' cellulosic ethanol enzyme kit: Enzymes for the hydrolysis of lignocellulosic materials. Bagsvaerd: Novozymes A/S.

Oboh, B. O., Ilori, M. O., Akinyemi, J. O., & Adebusoye, S. A. (2006). Hydrocarbon Degrading Potentials of Bacteria Isolated from a Nigerian Bitumen (Tarsand) Deposit. *Nature and Science*, *4*(3), 51–57.

Okano, K., Tanaka, T., Ogino, C., Fukuda, H., & Kondo, A. (2010). Biotechnological production of enantiomeric pure lactic acid from renewable resources: Recent achievements, perspectives, and limits. *Applied Microbiology and Biotechnology*, *85*(3), 413–423. doi:10.1007/s00253-009-2280-5

Oshima, T., & Imahori, K. (1974). Description of Thermus thermophilus (Yoshida and Oshima) comb. nov., a Nonsporulating Thermophilic Bacterium from a Japanese Thermal Spa. *International Journal of Systematic Bacteriology*, *24*(1), 102–112.

doi:10.1099/00207713-24-1-102

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., … Stevens, R. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research*, *42*(November 2013), 206–214. doi:10.1093/nar/gkt1226

Ozcan, B. D., Coskun, A., Ozcan, N., & Baylan, M. (2011). Some proteins of a new thermostable xylanase from alkaliphilic and thermophilic Bacillus sp. isolate DM-15. *Journal of Animal and Veterinary Advances*, *10*(2), 138–143.

Pal, A., & Khanum, F. (2010). Production and extraction optimization of xylanase from Aspergillus niger DFR-5 through solid-state-fermentation. *Bioresource Technology*, *101*(19), 7563–9. doi:10.1016/j.biortech.2010.04.033

Pan, P. G. (2006). Bioprospecting: issues and policy considerations. Retrieved from http://hawaii.gov/lrb/rpts06/bioconfs.html

Paper, J. M., Scott-Craig, J. S., Adhikari, N. D., Cuomo, C. a., & Walton, J. D. (2007). Comparative proteomics of extracellular proteins in vitro and in planta from the pathogenic fungus Fusarium graminearum. *Proteomics*, *7*(17), 3171–3183. doi:10.1002/pmic.200700184

Patskovsky, Y., Almo, S. C. (2009). Crystal structure of beta-1,4-xylosidase from Bacillus subtilis. Retrieved May 15, 2015, from http://www.rcsb.org/pdb/explore/explore.do?structureId=1yif

Payot, T., Chemaly, Z., & Fick, M. (1999). Lactic acid production by Bacillus coagulans — Kinetic studies and optimization of culture medium for batch and continuous fermentations, *0229*(98), 191–199.

Pedrosa, F. O., Monteiro, R. A., Wassem, R., Cruz, L. M., Ayub, R. a., Colauto, N. B., … Souza, E. M. (2011). Genome of herbaspirillum seropedicae strain SmR1, a specialized diazotrophic endophyte of tropical grasses. *PLoS Genetics*, *7*(5). doi:10.1371/journal.pgen.1002064

Peel, D., & Quayle, J. R. (1961). Microbial growth on C1 compounds. *Biochemical Journal*, *81*, 465–469.

Perego, M., Glaser, P., Minutello, A., Strauch, M. a., Leopold, K., & Fischer, W. (1995). Incorporation of D-alanine into lipoteichoic acid and wall teichoic acid in Bacillus subtilis: Identification of genes and regulation. *Journal of Biological Chemistry*. doi:10.1074/jbc.270.26.15598

Perez, C. M., Palmiano, E. P., Baun, L. C., & Juliano, B. O. (1971). Starch metabolism in the leaf sheaths and culm of rice. *Plant Physiology*, *47*(3), 404–8. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=365878&tool=pmcentrez&rendertype=abstract

Pérez-Avalos, O., Sánchez-Herrera, L. M., Salgado, L. M., & Ponce-Noyola, T. (2008). A bifunctional endoglucanase/endoxylanase from Cellulomonas flavigena with potential use in industrial processes at different pH. *Current Microbiology*, *57*(1), 39–44. doi:10.1007/s00284-008-9149-1

Perlin, a. S., & Suzuki, S. (1962). the Structure of Lichenin: Selective Enzymolysis Studies. *Canadian Journal of Chemistry*, *40*(1), 50–56. doi:10.1139/v62-009

Peterson, S. N., & Fraser, C. M. (2001). The complexity of simplicity. *Genome Biology*, *2*(2), COMMENT2002. doi:10.1016/S0378-4754(97)00044-X

Petti, C. a. (2007). Detection and identification of microorganisms by gene amplification and sequencing. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *44*(8), 1108–14. doi:10.1086/512818

Piao, H., Markillie, L. M., Culley, D. E., Mackie, R. I., & Hess, M. (2013). Improved Method for Isolation of Microbial RNA from Biofuel feedstock for metatranscriptomics. *Advances in Microbiology*, *3*, 101–107.

Pieper, R., Jha, R., Rossnagel, B., Van Kessel, A. G., Souffrant, W. B., & Leterme, P. (2008). Effect of barley and oat cultivars with different carbohydrate compositions on the intestinal bacterial communities in weaned piglets. *FEMS Microbiology Ecology*, *66*(3), 556–566. doi:10.1111/j.1574-6941.2008.00605.x

Pilli, S., Bhunia, P., Yan, S., LeBlanc, R. J., Tyagi, R. D., & Surampalli, R. Y. (2011). Ultrasonic pretreatment of sludge: a review. *Ultrasonics Sonochemistry*, *18*(1), 1–18. doi:10.1016/j.ultsonch.2010.02.014

Pinon, D., & Allen, C. (2013a). Syngenta Footprint for Enogen Corn Grows to 11 Ethanol Plants. Minnetonka: Syngenta. Retrieved from www.syngenta-us.com/News_releases/NewsPrint.aspx?id=174837

Pinon, D., & Allen, C. (2013b). Syngenta Pledges Support for Renewable Fuels. Minnetonka: Syngenta Crop Protection.

Plechkova, N. V, & Seddon, K. R. (2008). Applications of ionic liquids in the chemical industry. *Chemical Society Reviews*, *37*(1), 123–50. doi:10.1039/b006677j

Pointing, S. B. (1999). Qualitative methods for the determination of lignocellulolytic enzyme production by tropical fungi. *Fungal Diversity*, *2*(March), 17–33.

Polyprog. (2012). Polyurethane Manufacturer and Supplier. Retrieved from http://www.polyprog.co.uk/?gclid=CKe11KHpyrACFUYntAods2UzYA

Polz, M. F., & Cavanaugh, C. M. (1998). Bias in Template-to-Product Ratios in Multitemplate PCR. *Applied and Environmental Microbiology*, *64*(10), 3724–3730.

Proença, D. N., Francisco, R., Santos, C. V., Lopes, A., Fonseca, L., Abrantes, I. M. O., & Morais, P. V. (2010). Diversity of bacteria associated with Bursaphelenchus xylophilus and other nematodes isolated from Pinus pinaster trees with pine wilt disease. *PloS One*, *5*(12), e15191. doi:10.1371/journal.pone.0015191

Puchart, V., Vrsanská, M., Svoboda, P., Pohl, J., Ogel, Z. B., & Biely, P. (2004). Purification and characterization of two forms of endo-beta-1,4-mannanase from a thermotolerant fungus, Aspergillus fumigatus IMI 385708 (formerly Thermomyces lanuginosus IMI 158749). *Biochimica et Biophysica Acta*, *1674*(3), 239–50. doi:10.1016/j.bbagen.2004.06.022

Puls, J. (1997). Chemistry and biochemistry of hemicelluloses: relationship between hemicellulose structure and enzymes required for hyrdolysis. *Macromolecular Symposia*, *120*(1), 183–196.

Purdy, S. J., Maddison, A. L., Jones, L. E., Webster, R. J., Andralojc, J., Donnison, I., & Clifton-Brown, J. (2013). Characterization of chilling-shock responses in four genotypes of Miscanthus reveals the superior tolerance of M. x giganteus compared with M. sinensis and M. sacchariflorus. *Annals of Botany*, *111*(5), 999–1013. doi:10.1093/aob/mct059

Putnam, C. (2006). PROTEIN CALCULATOR v3.3. Retrieved March 10, 2015, from http://www.scripps.edu/~cdputnam/protcalc.html

Qi, M., Wang, P., O'Toole, N., Barboza, P. S., Ungerfeld, E., Leigh, M. B., … Forster, R. J. (2011). Snapshot of the eukaryotic gene expression in muskoxen rumen--a metatranscriptomic approach. *PloS One*, *6*(5), e20521. doi:10.1371/journal.pone.0020521

Qiagen. (2001). *QIAquick gel extraction protocol using a microcentrifuge*. Redwood City.

Qian, Y., Yomano, L. P., Preston, J. F., Aldrich, H. C., Ingram, O., & Ingram, L. O. (2003). Expression of the Klebsiella oxytoca Xylodextrin Utilization Operon (xynTB) in Escherichia coli Cloning , Characterization , and Functional Expression of the Klebsiella oxytoca Xylodextrin Utilization Operon (xynTB) in Escherichia coli †, *69*(10), 5957–5967. doi:10.1128/AEM.69.10.5957

Qureshi, N., Singh, V., Liu, S., Ezeji, T. C., Saha, B. C., & Cotta, M. a. (2014). Process integration for simultaneous saccharification, fermentation, and recovery (SSFR): Production of butanol from corn stover using Clostridium beijerinckii P260. *Bioresource Technology*, *154*, 222–228. doi:10.1016/j.biortech.2013.11.080

Raghothama, S., Simpson, P. J., Szabó, L., Nagy, T., Gilbert, H. J., & Williamson, M. P. (2000). Solution structure of the CBM10 cellulose binding module from Pseudomonas xylanase A. *Biochemistry*, *39*(5), 978–984. doi:10.1021/bi992163+

Rahman, K. S. M., Thahira-Rahman, J., Lakshmanaperumalsamy, P., & Banat, I. M. (2002). Towards efficient crude oil degradation by a mixed bacterial consortium. *Bioresource Technology*, *85*(3), 257–61. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12365493

Rajaram, S., & Varma, A. (1990). Production and characterization of xylanase from Bacillus thermoalkalophilus grown on agricultural wastes. *Applied Microbiology and Biotechnology*, *34*(1), 141–144. doi:10.1007/BF00170939

Rajoka, M. I., Bashir, A., & Malik, K. A. (1997). MUTAGENESIS OF CELLULOMONAS BIAZOTEA FOR ENHANCED PRODUCTION OF XYLANASES. *Bioresources Technology*, *62*, 99–108.

Ramesh, M. V., & Lonsane, B. K. (1989). Solid state fermentation for production of higher titres of thermostable alpha-amylase with two peaks for pH optima by Bacillus licheniformis M27. *Biotechnology Letters*, *11*(1), 49–52.

Rapp, P., & Wagner, F. (1986). Production and Properties of Xylan-Degrading Enzymes from Cellulomonas uda. *Applied an dEnvironmental Microbiology*, *51*(4), 746–752.

RCSB-PDB. (2005). Query Results: Bacillus coagulans. Retrieved May 15, 2015, from http://www.rcsb.org/pdb/results/results.do?qrid=122E13A8&tabtoshow=Current

Recchia, G. D., & Hall, R. M. (1997). Origins of the mobile gene cassettes found in integrons. *Trends In Microbiology*, *5*(10), 389–94. doi:10.1016/S0966-842X(97)01123-2

Reddy, G., Altaf, M., Naveena, B. J., Venkateshwar, M., & Kumar, E. V. (2008). Amylolytic bacterial lactic acid fermentation - a review. *Biotechnology Advances*, *26*(1), 22–34. doi:10.1016/j.biotechadv.2007.07.004

Regulapati, R., Malav, P. N., & Gummadi, S. N. (2007). Production of thermostable a-amylases by solid state fermentation - a review. *American Journal of Food Technology*, *2*(1), 1–11.

Reinhold-Hurek, B., & Hurek, T. (2011). Living inside plants: bacterial endophytes. *Current Opinion in Plant Biology*, *14*(4), 435–43. doi:10.1016/j.pbi.2011.04.004

Reis, V. M., Baldani, J. I., Baldani, V. L. D., & Dobereiner, J. (2000). Biological Dinitrogen Fixation in Gramineae and Palm Trees. *Critical Reviews in Plant Sciences*, *19*(3), 227–247. doi:10.1080/07352680091139213

Reisinger, V., & Eichacker, L. A. (2007). How to analyze protein complexes by 2D blue native SDS-PAGE. *Proteomics*, *7 Suppl 1*, 6–16. doi:10.1002/pmic.200700205

Rey, M. W., Ramaiya, P., Nelson, B. a, Brody-Karpin, S. D., Zaretsky, E. J., Tang, M., … Berka, R. M. (2004). Complete genome sequence of the industrial bacterium Bacillus licheniformis and comparisons with closely related Bacillus species. *Genome Biology*, *5*(10), R77. doi:10.1186/gb-2004-5-10-r77

Rhee, M. S., Moritz, B. E., Xie, G., Glavina Del Rio, T., Dalin, E., Tice, H., … Shanmugam, K. T. (2011). Complete Genome Sequence of a thermotolerant sporogenic lactic acid bacterium, Bacillus coagulans strain 36D1. *Standards in Genomic Sciences*, *5*(3), 331–40. doi:10.4056/sigs.2365342

Richardson, T. H., Tan, X., Frey, G., Callen, W., Cabell, M., Lam, D., … Miller, C. (2002). A novel, high performance enzyme for starch liquefaction. Discovery and optimization of a low pH, thermostable alpha-amylase. *The Journal of Biological Chemistry*, *277*(29), 26501–7. doi:10.1074/jbc.M203183200

Roberts, S. B., Gowen, C. M., Brooks, J. P., & Fong, S. S. (2010). Genome-scale metabolic analysis of Clostridium thermocellum for bioethanol production. *BMC Systems Biology*, *4*, 31. doi:10.1186/1752-0509-4-31

Roncato-Maccari, L. D. B., Ramos, H. J. O., Pedrosa, F. O., Alquini, Y., Chubatsu, L. S., Yates, M. G., … Souza, E. M. (2003). Endophytic Herbaspirillum seropedicae expresses nif genes in gramineous plants. *FEMS Microbiology Ecology*, *45*(1), 39–47. doi:10.1016/S0168-6496(03)00108-9

Roncero, M. I. (1983). Genes Controlling Xylan Utilization by Bacillus subtilis. *Journal of Bacteriology*, *156*(1), 257–263.

Rosegrant, M. W., Zhu, T., Msangi, S., & Sulser, T. (2008). Global Scenarios for Biofuels: Impacts and Implications. *Review of Agricultural Economics*, *30*(3), 495–505. doi:10.1111/j.1467-9353.2008.00424.x

Rosenblueth, M., & Martínez-Romero, E. (2006). Bacterial endophytes and their interactions with hosts. *Molecular Plant-Microbe Interactions : MPMI*, *19*(8), 827–37. doi:10.1094/MPMI-19-0827

Rothballer, M., Eckert, B., Schmid, M., Fekete, A., Schloter, M., Lehner, A., … Hartmann, A. (2008). Endophytic root colonization of gramineous plants by Herbaspirillum frisingense. *FEMS Microbiology Ecology*, *66*(1), 85–95. doi:10.1111/j.1574-6941.2008.00582.x

Rothschild, L. J., & Mancinelli, R. L. (2001). Life in extreme environments. *Nature*, *409*(6823), 1092–1101. doi:10.1038/35059215

Ruas-Madiedo, P., & de los Reyes-Gavilán, C. G. (2005). Invited review: methods for the screening, isolation, and characterization of exopolysaccharides produced by lactic acid bacteria. *Journal of Dairy Science*, *88*(3), 843–856. doi:10.3168/jds.S0022-0302(05)72750-8

Rubin, E. M. (2008). Genomics of cellulosic biofuels. *Nature*, *454*(7206), 841–5. doi:10.1038/nature07190

Ruller, R., Rosa, J. C., Faca, V. M., Greene, L. J., & Ward, R. J. (2006). Efficient constitutive expression of Bacillus subtilis xylanase A in Escherichia coli DH5alpha under the control of the Bacillus BsXA promoter. *Biotechnology and Applied Biochemistry*, *43*(Pt 1), 9–15. doi:10.1042/BA20050016

Saeidian, S., & Rashidzadeh, B. (2013). Optimum Temperature and Thermal Stability of Crude Polyphenol Oxidase In GreenSmall Cherry Tomatoe (Solanum Lycopersycum). *International Research Journal of Applied and Basic Sciences*, *4*(11), 3306–3311.

Sagaram, U. S., Deangelis, K. M., Trivedi, P., Andersen, G. L., Lu, S. E., & Wang, N. (2009). Bacterial diversity analysis of huanglongbing pathogen-infected citrus, using phyloChip arrays and 16S rRNA gene clone library sequencing. *Applied and Environmental Microbiology*, *75*(6), 1566–1574. doi:10.1128/AEM.02404-08

Saitou, N., & Nei, M. (1987). The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Molecular Biology and Evolution*, *4*(4), 406–425.

Sanchez, L. (2001). TCA protein precipitation protocol. *Www.imss.caltech.edu*. doi:10.1145/2003653.2003657

Sánchez-Herrera, L. M., Ramos-Valdivia, A. C., de la Torre, M., Salgado, L. M., & Ponce-Noyola, T. (2007). Differential expression of cellulases and xylanases by Cellulomonas flavigena grown on different carbon sources. *Applied Microbiology and Biotechnology*, *77*(3), 589–95. doi:10.1007/s00253-007-1190-7

Sánchez-Torres, J., Pérez, P., & Santamaría, R. I. (1996). A cellulase gene from a new alkalophilic Bacillus sp. (strain N186-1). Its cloning, nucleotide sequence and

expression in Escherichia coli. *Applied Microbiology and Biotechnology*, *46*(2), 149–155. doi:10.1007/s002530050797

Santos, C. R., Hoffmam, Z. B., de Matos Martins, V. P., Zanphorlin, L. M., de Paula Assis, L. H., Honorato, R. V., … Murakami, M. T. (2014). Molecular Mechanisms Associated with Xylan Degradation by Xanthomonas Plant Pathogens. *Journal of Biological Chemistry*, *289*(46), 32186–32200. doi:10.1074/jbc.M114.605105

Sasahara, K., Demura, M., & Nitta, K. (2002). Equilibrium and kinetic folding of hen egg-white lysozyme under acidic conditions. *Proteins*, *49*(4), 472–482. doi:10.1002/prot.10215

Sauer, M., Porro, D., Mattanovich, D., & Branduardi, P. (2008). Microbial production of organic acids: expanding the markets. *Trends in Biotechnology*, *26*(2), 100–8. doi:10.1016/j.tibtech.2007.11.006

Schäfer, G., Engelhard, M., & Müller, V. (1999). Bioenergetics of the Archaea. *Microbiology and Molecular Biology Reviews*, *63*(3), 570–620. doi:citeulike-article-id:11186509

Schallmey, M., Singh, A., & Ward, O. P. (2004). Developments in the use of Bacillus species for industrial production. *Canadian Journal of Microbiology*, *50*(1), 1–17. doi:10.1139/w03-076

Schönert, S., Seitz, S., Krafft, H., Feuerbaum, E. A., Andernach, I., Witz, G., & Dahl, M. K. (2006). Maltose and maltodextrin utilization by Bacillus subtilis. *Journal of Bacteriology*, *188*(11), 3911–3922. doi:10.1128/JB.00213-06

Schut, F., Vries, E. J. De, Gottschal, J. C., Betsy, R., Harder, W., Prins, R. A., & Button, D. K. (1993). Isolation of typical marine bacteria by dilution culture: growth, maintenance, and characteristics of isolates under laboratory conditions. *Applied and Environmental Microbiology*, *59*(7), 2150–2160.

Sedlak, M., & Ho, N. W. Y. (2004). Production of ethanol from cellulosic biomass hydrolysates using genetically engineered Saccharomyces yeast capable of cofermenting glucose and xylose. *Applied Biochemistry and Biotechnology*, *113-116*, 403–16. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15054267

Shah, S., & Gupta, M. N. (2008). The effect of ultrasonic pre-treatment on the catalytic activity of lipases in aqueous and non-aqueous media. *Chemistry Central Journal*, *2*, 1. doi:10.1186/1752-153X-2-1

Sharma, A., & Satyanarayana, T. (2011). Optimization of medium components and cultural variables for enhanced production of acidic high maltose-forming and Ca2+-independent α-amylase by Bacillus acidicola. *Journal of Bioscience and Bioengineering*, *111*(5), 550–3. doi:10.1016/j.jbiosc.2011.01.004

Shin, H. D., McClendon, S., Vo, T., & Chen, R. R. (2010). Escherichia coli binary culture engineered for direct fermentation of hemicellulose to a biofuel. *Applied and Environmental Microbiology*, *76*(24), 8150–8159. doi:10.1128/AEM.00908-10

Shrestha, P., Szaro, T. M., Bruns, T. D., & Taylor, J. W. (2011). Systematic search for cultivatable fungi that best deconstruct cell walls of Miscanthus and sugarcane in

the field. *Applied and Environmental Microbiology*, *77*(15), 5490–504. doi:10.1128/AEM.02996-10

Shulami, S., Zaide, G., Zolotnitsky, G., Langut, Y., Feld, G., Sonenshein, A. L., & Shoham, Y. (2007). A two-component system regulates the expression of an ABC transporter for xylo-oligosaccharides in Geobacillus stearothermophilus. *Applied and Environmental Microbiology*, *73*(3), 874–84. doi:10.1128/AEM.02367-06

Siegrist, J. (2013). Enzymatic food analysis. Retrieved from http://www.sigmaaldrich.com/technical-documents/articles/analytix/enzymatic-food-analysis.html

Siegumfeldt, H., Rechinger, K. B., & Jakobsen, M. (1999). Use of fluorescence ratio imaging for intracellular pH determination of individual bacterial cells in mixed cultures. *Microbiology (Reading, England)*, *145 (Pt 7*(1 999), 1703–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10439409

Simonen, M., & Palva, I. (1993). Protein secretion in Bacillus species. *Microbiological Reviews*, *57*(1), 109–37. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=372902&tool=pmcentrez&rendertype=abstract

Simpson, H. D., Haufler, U. R., & Daniel, R. M. (1991). An extremely thermostable xylanase from the thermophilic eubacterium Thermotoga. *The Biochemical Journal*, *277 (Pt 2*, 413–7. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1151249&tool=pmcentrez&rendertype=abstract

Šimůnek, J., Tishchenko, G., Rozhetsky, K., Bartoňová, H., Kopečný, J., & Hodrová, B. (2004). Chitinolytic enzymes from Clostridium aminovalericum: Activity screening and purification. *Folia Microbiologica*, *49*(2), 194–198. doi:10.1007/BF02931401

Sinitsyn, A. P., Gusakov, A. V., & Vlasenko, E. Y. (1991). Effect of Structural and Physico-Chemical Features of Cellulosic Substrates on the Efficiency of Enzymatic Hydrolysis. *Applied Biochemistry and Biotechnology*, *30*, 43–59.

Sivaramakrishnan, S., Gangadharan, D., Madhavan, K., Soccol, C. R., & Pandey, A. (2006). a-Amylases from microbial sources – an overview on recent developments. *Food Technology and Biotechnology*, *44*(2), 173–184.

So, C. M., & Young, L. Y. (1999). Isolation and Characterization of a Sulfate-Reducing Bacterium That Anaerobically Degrades Alkanes. *Applied and Environmental Microbiology*, *65*(7), 2969–2976.

Soemphol, W., Deeraksa, A., Matsutani, M., Yakushi, T., Toyama, H., Adachi, O., … Matsushita, K. (2011). Global Analysis of the Genes Involved in the Thermotolerance Mechanism of Thermotolerant Acetobacter tropicalis SKU1100. *Bioscience, Biotechnology, and Biochemistry*, *75*(10), 1921–1928. doi:10.1271/bbb.110310

Song, L. (2011). *STUDY AND ENGINEERING OF A GH11 ENDO-BETA-XYLANASE , A BIOMASS-DEGRADING HEMICELLULASE*. L'UNIVERSITE DE TOULOUSE.

Stampolidis, P., Kaderbhai, N. N., & Kaderbhai, M. a. (2009). Periplasmically-exported lupanine hydroxylase undergoes transition from soluble to functional inclusion bodies in Escherichia coli. *Archives of Biochemistry and Biophysics*, *484*(1), 8–15. doi:10.1016/j.abb.2009.01.017

Stephanopoulos, G. (2007). Challenges in engineering microbes for biofuels production. *Science (New York, N.Y.)*, *315*(5813), 801–4. doi:10.1126/science.1139612

Stephens, C., Christen, B., Watanabe, K., Fuchs, T., & Jenal, U. (2007). Regulation of D-xylose metabolism in Caulobacter crescentus by a LacI-type repressor. *Journal of Bacteriology*, *189*(24), 8828–34. doi:10.1128/JB.01342-07

Sticher, L., & Jones, R. L. (1992). Alpha-Amylase Isoforms are Posttranslationally Modified in the Endomembrane System of the Barley Aleurone Layer. *Plant Physiology*, *98*, 1080–1086. doi:10.1104/pp.98.3.1080

Stinson, M. W., & Merrick, J. M. (1974). Extracellular Enzyme Secretion by Pseudomonas lemoignei. *Journal of Bacteriology*, *119*(1), 152–161.

Stoker, H. S. (2012). Enzymes and Vitamins. In A. White & A. Landsberg (Eds.), *General, Organic and Biological Chemistry* (6th ed., pp. 414–457). Belmont: Brooks/Cole.

Stolp, H., & Starr, M. P. (1981). Principals of isolation, cultivation, and conservation of bacteria. In M. P. Starr, H. Stolp, H. G. Truper, A. Balows, & H. G. Schlegel (Eds.), *The Prokaryotes: A handbook on habitats, isolation and identification of bacteria* (1st ed., pp. 545–577). Berlin, Heidelberg: Springer-Verlag. doi:10.1007/978-3-642-30194-0

Straub, D., Yang, H., Liu, Y., Tsap, T., & Ludewig, U. (2013). Root ethylene signalling is involved in Miscanthus sinensis growth promotion by the bacterial endophyte Herbaspirillum frisingense GSF30 T. *Journal of Experimental Botany*, *64*(14), 4603–4615. doi:10.1093/jxb/ert276

Strobel, H. J., & Russell, J. B. (1986). Effect of pH and energy spilling on bacterial protein synthesis by carbohydrate-limited cultures of mixed rumen bacteria. *Journal of Dairy Science*, *69*(11), 2941–7. doi:10.3168/jds.S0022-0302(86)80750-0

Su, F., & Xu, P. (2014). Genomic analysis of thermophilic Bacillus coagulans strains: efficient producers for platform bio-chemicals. *Scientific Reports*, *4*, 3926. doi:10.1038/srep03926

Su, F., Yu, B., Sun, J., Ou, H.-Y., Zhao, B., Wang, L., … Xu, P. (2011). Genome sequence of the thermophilic strain Bacillus coagulans 2-6, an efficient producer of high-optical-purity L-lactic acid. *Journal of Bacteriology*, *193*(17), 4563–4. doi:10.1128/JB.05378-11

Subramaniyan, S., & Prema, P. (2002). Biotechnology of microbial xylanases: enzymology, molecular biology, and application. *Critical Reviews in Biotechnology*, *22*(1), 33–64. doi:10.1080/07388550290789450

Suga, K., Dedkm, G. van, & Moo-Young, M. (1975). Degradation of polysaccharides by endo and exo enzymes: a theoretical analysis. *Biotechnology and Bioengineering*,

17, 433–439.

Sukumaran, R. K. (2009). Bioethanol from lignocellulosic biomass: Part II production of cellulases and hemicellulases. In A. Pandey (Ed.), *Handbook of Plant Based Biofuels* (1st ed., Vol. 3, pp. 141–159). Boca Raton: CRC Press.

Sun, F. (2015). Why there are two bands in SDS-PAGE of a purified protein instead of one? Retrieved July 15, 2015, from https://www.researchgate.net/post/Why_there_are_two_bands_in_SDS-PAGE_of_a_purified_protein_instead_of_one2

Sun, Y., & Cheng, J. (2002). Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresource Technology*, *83*(1), 1–11. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12058826

Suslow, T. V., Schroth, M. N., & Isaka, M. (1982). Application of a Rapid Method for Gram Differentiation of Plant Pathogenic and Saprophytic Bacteria Without Staining. *Phytopathology*. doi:10.1094/Phyto-77-917

Suvd, D., Fujimoto, Z., Takase, K., Matsumura, M., & Mizuno, H. (2001). Crystal structure of Bacillus stearothermophilus alpha-amylase: possible factors determining the thermostability. *Journal of Biochemistry*, *129*(3), 461–468.

Suzuki, M. T., & Giovannoni, S. J. (1996). Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Applied and Environmental Microbiology*, *62*(2), 625–630.

Swatloski, R. P., Spear, S. K., Holbrey, J. D., & Rogers, R. D. (2002). Dissolution of Cellose with Ionic Liquids. *Journal of the American Chemical Society*, *124*(18), 4974–4975. doi:10.1021/ja025790m

Swiss Institute of Bioinformatics. (2014). Translate tool. Retrieved March 10, 2015, from http://web.expasy.org/translate/

Swiss Institute of Bioinformatics. (2015). ExPASy - Compute pI/Mw tool. Retrieved February 17, 2015, from http://web.expasy.org/compute_pi/

Syngenta. (2013a). Business Overview. In *Advisory Council Meeting* (pp. 1–10). Chicago: Syngenta. Retrieved from http://www.syngenta.com/country/us/en/agriculture/seeds/corn/enogen/stewardship/Documents/February 21, 2013/Business Overview.pdf

Syngenta. (2013b). Syngenta's corn amylase. Retrieved from http://www.syngenta.com/country/us/en/Enogen/Newsroom/Pages/SyngentasCornAmylase.aspx

Syngenta Seeds Inc. (2005). *Early Food Safety Evaluation* (Vol. 3272). Research Triangle Park. Retrieved from http://www.fda.gov/downloads/Food/FoodScienceResearch/Biotechnology/Submissions/UCM218880.pdf

Syvanen, M. (1994). Horizontal gene transfer: evidence and possible consequences. *Annual Review of Genetics*, *28*, 237–61. doi:10.1146/annurev.ge.28.120194.001321

Szila, A., Kardos, J., Osvath, S., Barna, L., & Zavodszky, P. (2007). Protein Folding. In A. Lajtha & N. Banik (Eds.), *Handbook of Neurochemistry and Molecular Neurobiology: Neural Protein Metabolism and Function* (3rd ed., p. 305–). Heidelberg: Springer-Verlag.

Tabil, L. G., Adapa, P. K., & Kashaninejad, M. (2011). Biomass Feedstock Pre-Processing - Part 1: Pre-Treatment. In M. A. Dos Santos Bernardes (Ed.), *Biofuel's Engineering Process Technology* (pp. 411–438). Rijeka: InTech.

Tamura, K., Nei, M., & Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(30), 11030–11035. doi:10.1073/pnas.0404206101

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, *30*(12), 2725–2729. doi:10.1093/molbev/mst197

Tarantino, L. M. (2013). *Agency Response Letter GRAS Notice No . GRN 000126*. Silver Spring. Retrieved from http://www.fda.gov/Food/IngredientsPackagingLabeling/GRAS/NoticeInventory/ucm153943.htm

Taylor, L. E., Weiner, R. M., Hutcheson, S. W., Ekborg, N. A., & Howard, M. (2014). Plant wall degradative compounds and systems. USA: United States Patent Application Publication. Retrieved from http://www.google.com/patents/US20140248688

Technische Universitat Braunschweig. (2014). Information on EC 3.2.1.8 - endo-1,4-beta-xylanase. Retrieved from http://www.brenda-enzymes.org/php/result_flat.php4?ecno=3.2.1.8

Teeri, T. T. (1996). Crystalline cellulose degradation: new insight into the function of cellobiohydrolases. *Trends in Biotechnology*, *15*, 160–167.

Tehei, M. (2005). Adaptation to extreme environments: Macromolecular dynamics in complex systems. *Biochimica et Biophysica Acta*, *1724*(3), 404–410.

Tellam, B. R., Winzor, D. J., & Nichol, L. W. (1978). The Role of Zinc in the Stabilization of the Dimeric Form of Bacterial a-Amylase. *Biochemical Journal*, *173*, 185–190.

Tenkanen, M., Makkonen, M., Perttula, M., Viikari, L., & Teleman, a. (1997). Action of Trichoderma reesei mannanase on galactoglucomannan in pine kraft pulp. *Journal of Biotechnology*, *57*(1-3), 191–204. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9335173

The Wales Centre of Excellence for Anaerobic Digestion. (2008). Mesophilic and Thermophilic Systems. Retrieved July 2, 2015, from http://www.walesadcentre.org.uk/technologies/mesophilicandthermophilicsystems.aspx

The Walser Group. (2012). Draft Genome of Pasteuria ramosa. Retrieved from http://evolution.unibas.ch/walser/pasteuria.htm

TheBioenergySite News Desk. (2011). BP Increases Biofuels Presence in Brazil. Retrieved from http://www.thebioenergysite.com/news/9562/bp-increases-biofuels-presence-in-brazil

TheBioenergySite News Desk. (2012). BP Biofuels Expand Operations at Brazil Facility. Retrieved from file:///C:/Users/lal8/Documents/documents/papers/industry/BP Biofuels Expand Operations at Brazil Facility - The Bioenergy Site.htm

Thermo Fisher Scientific Inc. (2014). CM0001, Nutrient Broth Product Detail. Retrieved June 10, 2015, from http://www.oxoid.com/UK/blue/prod_detail/prod_detail.asp?pr=CM0001&cat=&c=UK&lang=EN

ThermoWood. (2003). ThermoWood Handbook. Helsinki: Finnish ThermoWood Association.

Thompson, W. J. (1993). Cyclic nucleotide phosphodiesterases: Pharmocology, biochemistry and function. In C. W. Taylor (Ed.), *Intracellular Messengers* (1st ed., p. 289). Oxford: Pergamon Press.

TMO Renewables. (2012a). TMO Renewables takes part in Prime Minister's trade delegation to Brazil. Retrieved from http://www.tmo-group.com/tmo-renewables-takes-part-in-prime-ministers-trade-delegation-to-brazil/

TMO Renewables. (2012b). UK Prime Minister David Cameron endorses TMO's expertise and achievements during UKTI delegation to Brazil. Retrieved from http://www.tmo-group.com/uk-prime-minister-david-cameron-endorses-tmos-expertise-and-achievements-during-ukti-delegation-to-brazil/

TMO Renewables. (2013). Anglo-Brazilian Joint Venture to launch first 2G Commercial Cellulosic Ethanol Production Plant. Retrieved from http://www.tmo-group.com/anglo-brazilian-joint-venture-to-launch-first-2g-commercial-cellulosic-ethanol-production-plant/

Tobisch, S., Glaser, P., Krüger, S., & Hecker, M. (1997). Identification and characterization of a new beta-glucoside utilization system in Bacillus subtilis. *Journal of Bacteriology*, *179*(2), 496–506.

Toivari, M. H., Salusjärvi, L., Ruohonen, L., Salusja, L., & Penttila, M. (2004). Endogenous Xylose Pathway in Saccharomyces cerevisiae. *Applied and Environmental Microbiology*, *70*(6), 3681–3686. doi:10.1128/AEM.70.6.3681

Tomes, D., Lakshmanan, P., & Songstad, D. (2011). *Biofuels*. (D. Tomes, P. Lakshmanan, & D. Songstad, Eds.) (1st ed.). New York, NY: Springer New York. doi:10.1007/978-1-4419-7145-6

Torronen, A., & Rouvinen, J. (1995). Structural comparison of two major endo-1,4-xylanases from Trichoderma reesei. *Biochemistry*, *34*, 847–856.

Toth, E. M., Borsodi, A. K., Felfoldi, T., Vajna, B., Sipos, R., & Marialigeti, K. (2013). 7.4. Pheno- and genotypic characterisation of bacterial strains. In E. M. Toth & K. Marialigeti (Eds.), *Practical Microbiology* (1st ed., pp. 58–97). Budapest: Eötvös Loránd University. Retrieved from

http://elte.prompt.hu/sites/default/files/tananyagok/microbiology/index.html

Tse-Dinh, Y. C., Qi, H., & Menzel, R. (1997). DNA supercoiling and bacterial adaptation: Thermotolerance and thermoresistance. *Trends in Microbiology*, *5*(8), 323–326. doi:10.1016/S0966-842X(97)01080-9

Turner, P., Mamo, G., & Karlsson, E. N. (2007a). Additional file 1. *Microbial Cell Factories*, *6*(3), 226–228. doi:10.1186/1475-2859-6-9

Turner, P., Mamo, G., & Karlsson, E. N. (2007b). Additional file 2. *Microbial Cell Factories*, *6*, 9. doi:10.1186/1475-2859-6-9

Turner, P., Mamo, G., & Karlsson, E. N. (2007c). Potential and utilization of thermophiles and thermostable enzymes in biorefining. *Microbial Cell Factories*, *6*, 9. doi:10.1186/1475-2859-6-9

Uffen, R. L. (1997). Xylan degradation: a glimpse at microbial diversity. *Journal of Industrial Microbiology & Biotechnology*, *19*, 1–6.

UniProt Consortium. (2014). Xylose isomerase - xylA - Thermoanaerobacterium thermosulfurigenes. Retrieved from http://www.uniprot.org/uniprot/P19148

Upadhyay, M. K., Sharma, R., Pandey, A. K., & Rajak, R. A. M. C. (2005). An improved zymographic method for detection of amylolytic enzymes of fungi on polyacrylamide gels. *Mycologist*, *19*(November), 138–140. doi:10.1017/S0269915XO5004015

van Aalten, D. M., Komander, D., Synstad, B., Gåseidnes, S., Peter, M. G., & Eijsink, V. G. (2001). Structural insights into the catalytic mechanism of a family 18 exo-chitinase. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(16), 8979–8984. doi:10.1073/pnas.151103798

van den Burg, B. (2003). Extremophiles as a source for novel enzymes. *Current Opinion in Microbiology*, *6*(3), 213–218. doi:10.1016/S1369-5274(03)00060-2

van der Veen, B. E., Harris, H. M., O´Toole, P. W., & Claesson, M. J. (2014). Metaphor: Finding Bi-directional Best Hit homology relationships in (meta)genomic datasets. *Genomics*, *104*(6), 459–463. doi:10.1016/j.ygeno.2014.10.008

van Dijl, J. M., & Hecker, M. (2013). Bacillus subtilis: from soil bacterium to super-secreting cell factory. *Microbial Cell Factories*, *12*, 3. doi:10.1186/1475-2859-12-3

van Dyk, J. S., Sakka, M., Sakka, K., & Pletschke, B. I. (2009). The cellulolytic and hemi-cellulolytic system of Bacillus licheniformis SVD1 and the evidence for production of a large multi-enzyme complex. *Enzyme and Microbial Technology*, *45*(5), 372–378. doi:10.1016/j.enzmictec.2009.06.016

van Dyk, J. S., Sakka, M., Sakka, K., & Pletschke, B. I. (2010). Characterisation of the multi-enzyme complex xylanase activity from Bacillus licheniformis SVD1. *Enzyme and Microbial Technology*, *47*(4), 174–177. doi:10.1016/j.enzmictec.2010.06.004

Van Pelt, C., Verduin, C. M., Goessens, W. H. F., Vos, M. C., Tümmler, B., Segonds, C., … Van Belkum, A. (1999). Identification of Burkholderia spp. in the clinical microbiology laboratory: Comparison of conventional and molecular methods.

Journal of Clinical Microbiology, *37*(7), 2158–2164.

Van Soestbergen, A. A., & Lee, C. H. (1969). Pour plates or streak plates? *Applied Microbiology*, *18*(6), 1092–3. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=378201&tool=pmcentrez&rendertype=abstract

Vaneechoutte, M., Dijkshoorn, L., Nemec, A., Kämpfer, P., & Wauters, G. (2011). Acinetobacter, Achromobacter, Chryseobacterium, Moraxella, and other nonfermentative gram-negative rods. In M. L. L. and D. W. W. James Versalovic, Karen C Carroll, Guido Funke, James H Jorgensen (Ed.), *Manual of clinical microbiology* (pp. 714–738). ASM Press. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Acinetobacter,+Chryseobacterium,+Moraxella,+and+Other+Nonfermentative+Gram-Negative+Rods#0

VanFossen, A. L., Verhaart, M. R. a, Kengen, S. M. W., & Kelly, R. M. (2009). Carbohydrate utilization patterns for the extremely thermophilic bacterium Caldicellulosiruptor saccharolyticus reveal broad growth substrate preferences. *Applied and Environmental Microbiology*, *75*(24), 7718–7724. doi:10.1128/AEM.01959-09

Varrot, A., Schulein, M., & Davies, G. J. (1999). Structural Changes of the Active Site Tunnel of Humicola insolens cellobiohydrolase, Cel6A, upon oligosaccharide binding. *Biochemistry*, *38*, 8884–8891.

Villegas, J. D., & Gnansounou, E. (2008). Techno-economic and environmental evaluation of lignocellulosic biochemical refineries: need for a modular platform for integrated assessment (MPIA). *Journal of Scientific and Industrial Research*, *67*, 927–940. Retrieved from http://infoscience.epfl.ch/record/124978

Vintila, T., Dragomirescu, M., Croitoriu, V., Vintila, C., Barbu, H., & Sand, C. (2010). Saccharification of lignocellulose - with reference to Miscanthus - using different cellulases. *Romanian Biotechnological Letters*, *15*(4), 5498–5504.

Vrije, T. De, Haas, G. G. De, Tan, G. B., Keijsers, E. R. P., & Claassen, P. A. M. (2002). Pretreatment of Miscanthus for hydrogen production by Thermotoga elfii. *International Journal of Hydrogen Energy*, *27*, 1381–1390.

Walfridsson, M., Bao, X., Anderlund, M., Lilius, G., Bülow, L., & Hahn-Hägerdal, B. (1996). Ethanolic fermentation of xylose with Saccharomyces cerevisiae harboring the Thermus thermophilus xylA gene, which expresses an active xylose (glucose) isomerase. *Applied and Environmental Microbiology*, *62*(12), 4648–51. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=168291&tool=pmcentrez&rendertype=abstract

Walker, J. C. F., Butterfield, B. G., Langrish, T. A. G., Harris, J. M., & Uprichard, J. M. (1993). Basic wood chemistry and cell wall ultrastructure. In J. C. F. Walker (Ed.), *Primary Wood Processing Principles and Practice* (First., pp. 23–67). Oxford: Chapman & Hall.

Walsum, G. P., Allen, S., Spencer, M., Laser, M., Antal, M., & Lynd, L. (1996). Conversion of lignocellulosics pretreated with liquid hot water to ethanol. *Applied Biochemistry and Biotechnology*, *57-58*(1), 157–170. doi:10.1007/BF02941696

Wan, C., & Li, Y. (2011). Effect of hot water extraction and liquid hot water pretreatment on the fungal degradation of biomass feedstocks. *Bioresource Technology*, *102*(20), 9788–93. doi:10.1016/j.biortech.2011.08.004

Wang, B., Wang, X., & Feng, H. (2010). Deconstructing recalcitrant Miscanthus with alkaline peroxide and electrolyzed water. *Bioresource Technology*, *101*(2), 752–60. doi:10.1016/j.biortech.2009.08.063

Wang, L., Wang, G., Li, S., & Jiang, J. (2011). Luteibacter jiangsuensis sp. nov.: a methamidophos-degrading bacterium isolated from a methamidophos-manufacturing factory. *Current Microbiology*, *62*(1), 289–95. doi:10.1007/s00284-010-9707-1

Wang, L., Zhao, B., Liu, B., Yu, B., Ma, C., Su, F., … Xu, P. (2010). Efficient production of L-lactic acid from corncob molasses, a waste by-product in xylitol production, by a newly isolated xylose utilizing Bacillus sp. strain. *Bioresource Technology*, *101*(20), 7908–7915. doi:10.1016/j.biortech.2010.05.031

Wang, S. L., Liang, Y. C., & Liang, T. W. (2011). Purification and characterization of a novel alkali-stable a-amylase from Chryseobacterium taeanense TKU001, and application in antioxidant and prebiotic. *Process Biochemistry*, *46*(3), 745–750. doi:10.1016/j.procbio.2010.11.022

Wass, M. N., Kelley, L. a, & Sternberg, M. J. E. (2010). 3DLigandSite: predicting ligand-binding sites using similar structures. *Nucleic Acids Research*, *38*(Web Server issue), W469–73. doi:10.1093/nar/gkq406

Watanabe, K., Hata, Y., Kizaki, H., Katsube, Y., & Suzuki, Y. (1997). The refined crystal structure of Bacillus cereus oligo-1,6-glucosidase at 2.0 A resolution: structural characterization of proline-substitution sites for protein thermostabilization. *Journal of Molecular Biology*, *269*(1), 142–153. doi:10.1006/jmbi.1997.1018

Werpy, T., & Petersen, G. (2004). *Top Value Added Chemicals from Biomass Volume I — Results of Screening for Potential Candidates from Sugars and Synthesis Gas*. Oak Ridge. Retrieved from http://www.osti.gov/bridge

Westereng, B., Ishida, T., Vaaje-Kolstad, G., Wu, M., Eijsink, V. G. H., Igarashi, K., … Sandgren, M. (2011). The putative endoglucanase PcGH61D from Phanerochaete chrysosporium is a metal-dependent oxidative enzyme that cleaves cellulose. *PloS One*, *6*(11), e27807. doi:10.1371/journal.pone.0027807

Wheatley, M. a, & Moo-Young, M. (1977, February). Degradation of polysaccharides by endo- and exoenzymes: dextran-dextranase model systems. *Biotechnology and Bioengineering*. doi:10.1002/bit.260190206

Williams, A. G. (1983). Staining reactions for the detection of hemicellulose-degrading bacteria. *FEMS Microbiology Letters*, *20*(2), 253–258. doi:http://dx.doi.org/

Winterhalter, C., & Liebl, W. (1995). Two Extremely Thermostable Xylanases of the

Hyperthermophilic Bacterium Thermotoga MSB8. *Applied and Environmental Microbiology*, *61*(5), 1810–1815.

Wittig, I., Braun, H.-P., & Schagger, H. (2006). Blue native PAGE. *Nature Protocols*, *1*(1), 418–428. Retrieved from http://dx.doi.org/10.1038/nprot.2006.62

Wittig, I., & Schägger, H. (2005). Advantages and limitations of clear-native PAGE. *Proteomics*, *5*(17), 4338–46. doi:10.1002/pmic.200500081

Wong, K. K., Tan, L. U., & Saddler, J. N. (1988). Multiplicity of beta-1,4-xylanase in microorganisms: functions and applications. *Microbiological Reviews*, *52*(3), 305–17. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=373146&tool=pmcentrez&rendertype=abstract

Wooley, R., Ruth, M., Sheehan, J., Ibsen, K., Majdeski, H., & Galvez, A. (1999). *Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis Current and Futuristic Scenarios*. Colorado.

Wu, J. H. D., & Newcomb, M. (2009). Genetic Engineered Clostridium Thermocellum for Ethanol Production from Cellulose. Retrieved from http://rochester.technologypublisher.com/technology/2966

Wyman, C. E., & Yang, B. (2009). Cellulosic biomass could help meet California's transportation fuel needs. *California Agriculture*, *63*(4), 185–190.

Xiao, Z., Storms, R., & Tsang, A. (2006a). A quantitative starch-iodine method for measuring alpha-amylase and glucoamylase activities. *Analytical Biochemistry*, *351*(1), 146–8. doi:10.1016/j.ab.2006.01.036

Xiao, Z., Storms, R., & Tsang, A. (2006b). A quantitative starch-iodine method for measuring alpha-amylase and glucoamylase activities. *Analytical Biochemistry*, *351*(1), 146–8. doi:10.1016/j.ab.2006.01.036

Xu, C., Huang, R., Teng, L., Wang, D., Hemme, C. L., Borovok, I., … Xu, J. (2013). Structure and regulation of the cellulose degradome in Clostridium cellulolyticum. *Biotechnology for Biofuels*, *6*(1), 73. doi:10.1186/1754-6834-6-73

Xu, Z. H., Bai, Y. L., Xu, X., Shi, J. S., & Tao, W. Y. (2005). Production of alkali-tolerant cellulase-free xylanase by Pseudomonas sp. WLUN024 with wheat bran as the main substrate. *World Journal of Microbiology and Biotechnology*, *21*(4), 575–581. doi:10.1007/s11274-004-3491-7

Yáñez, R., Moldes, A. B., Alonso, J. L., & Parajó, J. C. (2003). Production of D(-)-lactic acid from cellulose by simultaneous saccharification and fermentation using Lactobacillus coryniformis subsp. torquens. *Biotechnology Letters*, *25*(14), 1161–1164. doi:10.1023/A:1024534106483

Yang, B., Willies, D. M., & Wyman, C. E. (2006). Changes in the Enzymatic Hydrolysis Rate of Avicel Cellulose With Conversion. *Biotechnology and Bioengineering*, *94*(6), 1122–1128. doi:10.1002/bit

Yang, B., & Wyman, C. E. (2006). BSA Treatment to Enhance Enzymatic Hydrolysis of

Cellulose in Lignin Containing Substrates. *Biotechnology and Bioengineering*, *94*(4), 611–617. doi:10.1002/bit

Yang, H., Yan, R., Chen, H., Lee, D. H., & Zheng, C. (2007). Characteristics of hemicellulose, cellulose and lignin pyrolysis. *Fuel*, *86*(12-13), 1781–1788. doi:10.1016/j.fuel.2006.12.013

Yang, Y., Zhang, W., Huang, J., Lin, L., Lian, H., Lu, Y., … Wang, S. (2010). Purification and characterization of an extracellular xylanase from Aspergillus niger C3486. *African Journal Of Microbiology Research*, *4*(21), 2249–2256.

Yarbrough, J. M., Himmel, M. E., & Ding, S.-Y. (2009). Plant cell wall characterization using scanning probe microscopy techniques. *Biotechnology for Biofuels*, *2*, 17. doi:10.1186/1754-6834-2-17

Yoshida, M., Liu, Y., Uchida, S., Kawarada, K., Ukagami, Y., Ichinose, H., … Fukuda, K. (2008). Effects of cellulose crystallinity, hemicellulose, and lignin on the enzymatic hydrolysis of Miscanthus sinensis to monosaccharides. *Bioscience, Biotechnology, and Biochemistry*, *72*(3), 805–810. doi:10.1271/bbb.70689

Yu, F. B., Shan, S. D., Luo, L. P., Guan, L. B., & Qin, H. (2013). Isolation and characterization of a Sphingomonas sp. strain F-7 degrading fenvalerate and its use in bioremediation of contaminated soil. *Journal of Environmental Science and Health. Part. B, Pesticides, Food Contaminants, and Agricultural Wastes*, *48*(3), 198–207. doi:10.1080/03601234.2013.730299

Yu, L., Xu, M., Tang, I.-C., & Yang, S.-T. (2015). Metabolic engineering of Clostridium tyrobutyricum for n-butanol production through co-utilization of glucose and xylose. *Biotechnology and Bioengineering*, *9999*(xxx), n/a–n/a. doi:10.1002/bit.25613

Yun, Y.-M., Jung, K.-W., Kim, D.-H., Oh, Y.-K., Cho, S.-K., & Shin, H.-S. (2013). Optimization of dark fermentative H_2 production from microalgal biomass by combined (acid + ultrasonic) pretreatment. *Bioresource Technology*, *141*, 220–6. doi:10.1016/j.biortech.2013.02.054

Zeeman, S. C., Kossmann, J., & Smith, A. M. (2010). Starch: its metabolism, evolution, and biotechnological modification in plants. *Annual Review of Plant Biology*, *61*, 209–34. doi:10.1146/annurev-arplant-042809-112301

Zhang, F., Rodriguez, S., & Keasling, J. D. (2011). Metabolic engineering of microbial pathways for advanced biofuels production. *Current Opinion in Biotechnology*, *22*(6), 775–783. doi:10.1016/j.copbio.2011.04.024

Zhang, J., Li, J., Chen, L., & Thring, R. W. (2011). Remediation of Refinery Oily Sludge Using Isolated Strain and Biosurfactant. In *2011 International Symposium on Water Resource and Environmental Protection (ISWREP) (Volume:3)* (pp. 1649 – 1653). Xi'an: IEEE. doi:10.1109/ISWREP.2011.5893355

Zhang, M., Su, R., Qi, W., & He, Z. (2010). Enhanced enzymatic hydrolysis of lignocellulose by optimizing enzyme complexes. *Applied Biochemistry and Biotechnology*, *160*(5), 1407–14. doi:10.1007/s12010-009-8602-3

Zhang, X., Jantama, K., Moore, J. C., Jarboe, L. R., Shanmugam, K. T., & Ingram, L. O. (2009). Metabolic evolution of energy-conserving pathways for succinate production in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(48), 20180–20185. doi:10.1073/pnas.0905396106

Zhang, X., Tu, M., & Paice, M. G. (2011). Routes to Potential Bioproducts from Lignocellulosic Biomass Lignin and Hemicelluloses. *BioEnergy Research*, *4*(4), 246–257. doi:10.1007/s12155-011-9147-1

Zhang, Y. H. P., Sun, J., & Zhong, J. J. (2010). Biofuel production by in vitro synthetic enzymatic pathway biotransformation. *Current Opinion in Biotechnology*, *21*(5), 663–669. doi:10.1016/j.copbio.2010.05.005

Zhang, Z., Song, Y., Jiang, K., Xue, Y., & Ma, Y. (2015). Characterization of solute-binding protein XynE of the xylooligosaccharide transporter from Bacillus sp. N16-5. *Wei sheng wu xue bao = Acta microbiologica Sinica*, *55*(1), 40–49.

Zhao, Y., Haney, M. J., Klyachko, N. L., Li, S., Booth, S. L., Higginbotham, S. M., … Batrakova, E. V. (2011). Polyelectrolyte complex optimization for macrophage delivery of redox enzyme nanoparticles. *Nanomedicine (London, England)*, *6*(1), 25–42. doi:10.2217/nnm.10.129

Zúñiga, M., Pérez, G., & González-Candelas, F. (2002). Evolution of arginine deiminase (ADI) pathway genes. *Molecular Phylogenetics and Evolution*, *25*(3), 429–444. doi:10.1016/S1055-7903(02)00277-4

Zverlov, V., Mahr, I. S., Riedel, K., & Bronnenmeier, K. (1998). Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from the extreme with separate glycosyl hydrolase family 9 and 48 catalytic domains. *Microbiology*, *144*, 457–465.

Chapter 8 Appendices

8.1 Appendix 1: Isolation and Characterisation of Bacterial Library

8.1.1 Bacteria associated with *Miscanthus*

Table 8.1 Bacteria isolated from or associated with *Miscanthus* and related feedstocks that were reported in the literature.

Feedstock	Microbe	Reference	Application	Reference
Miscanthus x giganteus	*Acidocella sp.*	(D. Li, 2011)	Acidophilic	(Jones, Hedrich, & Johnson, 2013; D. Li, 2011)
	Actinomycetes	(Klamer & Bååth, 1998)	Commercial drug industry Thermophilic	(McNeil & Brown, 1994)
	Azospirillum doebereinerae sp. nov.	(Eckert *et al.*, 2001)	Nitrogen fixing	(Eckert *et al.*, 2001)
	Azospirillum lipoferum	(Eckert *et al.*, 2001; Kirchhof *et al.*, 1997)	Nitrogen fixing	(Eckert *et al.*, 2001; Kirchhof *et al.*, 1997)
	Azospirillum sp.	(Eckert *et al.*, 2001)	Nitrogen fixing	(Eckert *et al.*, 2001; Hartmann *et al.*, 1995)
	Bacillus sp.	(Eiland, Klamer, Lind, Leth, & Bååth, 2001; Klamer & Bååth, 1998)	Industrial enzyme production Biotechnology	(A. de Boer *et al.*, 1994; L. Liu *et al.*, 2013) (Schallmey *et al.*, 2004)
	Burkholderia sp.	(D. Li, 2011)	Commercial drug industry Nitrogen fixing Human pathogen Biotechnology	(McNeil & Brown, 1994) (Mohanta *et al.*, 2010; Reis, Baldani, Baldani, & Dobereiner, 2000) (Coenye *et al.*, 2001; Du Toit, 2015; Van Pelt *et al.*, 1999)(C. H. Liu *et al.*, 2006; Mohana, Shah, Divecha, & Madamwar, 2008)
	Desulfovibrio sp.	(D. Li, 2011)	Metal reducing Human pathogen	(Callaghan, Wawrik, Ní Chadhain, Young, & Zylstra, 2008; So & Young, 1999) (Goldstein, Citron, Peraino, & Cross, 2003)
	Fulvimonas sp.	(D. Li, 2011)	Biotechnology	(Mergaert *et al.*, 2002)
	Geobacter sp.	(D. Li, 2011)	Metal reducing	(Caccavo *et al.*, 1994)
	Methyloversatilis sp.	(D. Li, 2011)	Thermophilic Biotechnology	(Doronina, Kaparullina, & Trotsenko, 2014) (Cai, Qian, Cai, & Chen, 2011)
	Oxalicibacterium sp.	(D. Li, 2011)		
	Pseudomonas sp.	(D. Li, 2011)	Biotechnology	(Kordel, Hofmann, Schomburg, & Schmid, 1991;

				Raghothama *et al.*, 2000; Z. H. Xu *et al.*, 2005)
	Raoultella sp.	(D. Li, 2011)	Biotechnology	(Kimura, Chung, Itoh, Hiraishi, & Okabe, 2014)
	Singulisphaera sp.	(D. Li, 2011)	Acidophilic	(Kulichevskaya *et al.*, 2008, 2011; D. Li, 2011)
	Sphingomonas sp.	(D. Li, 2011)	Biotechnology	(F. B. Yu *et al.*, 2013)
	Herbaspirillum frisingense	(Kirchhof *et al.*, 2001)	Nitrogen fixing	(Pedrosa *et al.*, 2011; Roncato-Maccari *et al.*, 2003; Rothballer *et al.*, 2008)
	Herbaspirillum seropedicae	(Kirchhof *et al.*, 1997)	Nitrogen fixing	(Hartmann *et al.*, 1995; Reis *et al.*, 2000)
Miscanthus sinensis	*Clostridium aminovalericum*	(Miyamoto, Kawahara, & Minamisawa, 2004; Rosenblueth & Martínez-Romero, 2006)	Biotechnology	(Šimůnek *et al.*, 2004)
	Clostridium intestinale	(Miyamoto *et al.*, 2004)	Human pathogen	(Elsayed & Zhang, 2005)
	Clostridium acetobutylicum	(Miyamoto *et al.*, 2004)	Biotechnology	(Lütke-Eversloh & Bahl, 2011)
	Clostridium beijerincki	(Miyamoto *et al.*, 2004)	Biotechnology	(T. Guo *et al.*, 2012; Qureshi *et al.*, 2014)
Miscanthus spp.	*Herbaspirillum frisingense*	(Reis *et al.*, 2000)	Nitrogen fixation	(Straub, Yang, Liu, Tsap, & Ludewig, 2013)

8.1.2 Morphology of the bacterial library

Table 8.2 Identification table of bacterial isolates.

Yellow – mixed colony that could not be separated; red –does not match 16S identification (ID); blue – contaminated or BLAST hit below 97 %. ND – not determined; RT – room temperature; o/n – overnight; * - cultivated on MRS media; ^ - cultivated on Wilkins-Charlgren media; all other isolates cultivated on Nutrient Agar medium.

Clade	16S ID	Gram	Shape	Isolation temperature	Days until growth	Colour	Surface	Form	Elevation	Margin	Transparency
1	*Burkholdeia fungorum* ABEb	mixed	cocci/cigars	RT	5	cream	shiny	punctiform	raised	entire	opaque/translucent edge
1	*Burkholdeia fungorum* ABF1a	mixed	rods	RT	5	cream	dry	punctiform	raised	undulate	opaque
1	*Burkholdeia fungorum* ABF1b	mixed	rods	RT	5	cream	shiny	punctiform	flat	entire	translucent
1	*Burkholderia sordidicola* ADC	positive	rod	RT	7	ND	ND	ND	ND	ND	ND
1	*Burkholderia sp.* ADJa	negative	cigar	RT	5	ND	ND	ND	ND	ND	ND
1	*Burkholderia sp.* ADJb	negative	cigar	RT	5	ND	ND	ND	ND	ND	ND
1	*Burkholderia fungorum* ADR	positive	rod	RT	7	yellow	shiny	circular	flat	entire	translucent
1	*Burkholdeia fungorum* AEIa	mixed	rods	RT	5	white	dry	punctiform	raised	undulate	opaque
1	*Burkholdeia fungorum* AEIb	mixed	rods	RT	5	white	dry	punctiform	raised	undulate	opaque
1	*Burkholderia fungorum* AELa	positive	rod	RT	11	white	shiny	circular	flat	entire	translucent
1	*Burkholderia fungorum* AELb	negative/ variable	rod	RT	11	white	shiny	circular	flat	entire	translucent
1	*Burkholderia sp.* AEM	negative	cigar/rod	RT	8	ND	ND	ND	ND	ND	ND
1	*Burkholderia sordidicola* ADV	negative	cocci	RT	7	yellow	shiny	punctiform	raised	entire	opaque
1	*Burkholderia sordidicola* ADX 1	positive	cigar	RT	7	cream	shiny	punctiform	flat	entire	opaque
1	*Burkholderia fungorum*	positive	rod/cigar	RT	8	grey	shiny	circular	flat	entire	opaque

	AES 1					white					
1	*Burkholdeia fungorum* AEW2a	mixed	cigars/rods	RT	5	cream	shiny	circular	flat	entire	translucent
1	*Burkholdeia fungorum* AEW2b	mixed	cigars/rods	RT	5	yellow	shiny	circular	flat	entire	translucent
1	*Burkholderia fungorum* AEXa	positive	cigar	RT	5	white	shiny	punctiform	flat	entire	opaque/translucent edge
1	*Burkholderia fungorum* AEXb	variable	cigar	RT	5	white	shiny	punctiform	flat	entire	opaque/translucent edge
1	*Burkholdeia fungorum* AEYa	mixed	cigars/rods	RT	5	cream	shiny	circular	flat	entire	translucent
1	*Burkholdeia fungorum* AEYb	mixed	cigars/rods	RT	5	cream	shiny	circular	flat	entire	translucent
1	*Burkholderia fungorum* AFA	negative	cocci	RT	8	white	shiny	circular	raised	entire/ filamentous	opaque
1	*Burkholderia fungorum* AFC	positive	rod	RT	8	white	shiny	punctiform	flat	entire/erose	opaque
1	*Burkholderia fungorum* AFDa	mixed	rods/cocci	RT	5	cream	dry	circular	flat	entire	opaque/translucent edge
1	*Burkholderia fungorum/L. rhizovicinus* AFDb	negative	rods	RT	5	cream	shiny	circular	flat	entire	translucent
1	*Burkholdeia fungorum* AFEb	mixed	cocci/cigar	RT	5	cream	shiny	circular	flat	entire	translucent
2	*Enhydrobacter sp. Or Moraxella osloensis* ABF2	positive	cocci/cigar	RT	3	white	shiny	punctiform	raised	entire	opaque
2	*Enhydrobacter aerosaccus/Moraxella sp.* ACB	negative	rod	RT	3	cream	shiny	punctiform	crateriform	entire	opaque
2	*Pedobacter terrae* ACL	positive	rod	RT	3	pink	shiny	circular	flat	entire	translucent
2	*Enhydrobacter sp./Moraxella sp.* ADK	variable	cocci	RT	7	white	shiny	punctiform	flat	entire	translucent
2	*Micrococcus luteus* ADM	negative	cocci	RT	7	yellow	dry	punctiform	convex	entire	opaque
2	*Microbacterium sp.* ADQ	positive	rod	RT	7	yellow	shiny	circular	raised	entire	opaque
2	*Microbacterium sp.* ADS	positive	cigar/rod	RT	7	yellow	shiny	punctiform	flat	entire	translucent
2	*Phyllobacterium ifriqiyense* ADT	positive	rod	RT	7	white	shiny	punctiform	raised	entire	opaque
2	*Massilia dura* ADW	negative	rod	RT	7	ND	ND	ND	ND	ND	ND

2	*Microbacterium sp.* AEP	positive	cocci	RT	8	yellow	shiny	punctiform	flat	entire	translucent
2	*Herbaspirillum hiltneri* AEQ	positive	rod	RT	8	cream	shiny	punctiform	flat	entire	translucent
2	*Microbacterium sp.* AEU	positive	cocci	RT	8	yellow	shiny	circular	flat	entire	translucent
2	*Dyella sp.* AFG	positive	rod	RT	8	yellow	shiny	punctiform	raised	entire	translucent
3	*Luteibacter sp.* ACN	positive	rod	RT	3	yellow	shiny	circular	flat	erose	translucent
3	*Luteibacter rhizovicinus* ADB	negative	rod	RT	7	yellow	shiny	punctiform	raised	entire	translucent
3	*Lutiebacter rhizovicinus* ADF	positive	rod	RT	7	yellow	shiny	punctiform	flat	entire	translucent
3	*Luteibacter rhizovicinus* ADN	variable	rod	RT	7	yellow	shiny	circular	raised	entire	translucent
3	*Luteibacter sp.* ADP	positive	rod	RT	7	yellow	shiny	circular	flat	entire	translucent
3	*Luteibacter rhizovicinus* ADX 2	positive	rod	RT	7	yellow	shiny	punctiform	raised	entire	translucent
3	*Luteibacter rhizovicinus* AEB (M4)	positive	rod	RT	8	yellow	shiny	punctiform	flat	entire	translucent
3	*Luteibacter sp.* AEC	positive	rod	RT	8	yellow	shiny	punctiform	flat	entire	translucent
3	*Luteibacter sp.* AEK	positive	rod	RT	8	yellow	shiny	circular	raised	entire	translucent
3	*Luteibacter rhizovicinus* AEOa	mixed	rods/cigars	RT	5	cream	shiny	punctiform	raised	entire	translucent
3	*Luteibacter rhizovicinus* AEOb	mixed	cigars/cocci	RT	5	yellow	shiny	punctiform	raised	erose	opaque
3	*Luteibacter rhizovicinus* AER	positive	rod	RT	8	yellow	shiny	punctiform	raised	entire	translucent
3	*Luteibacter rhizovicinus* AES 2	positive	rod	RT	8	yellow	shiny	punctiform	raised	entire	opaque
3	*Luteibacter rhizovicinus* AEV	positive	rod/cigar	RT	8	yellow	shiny	circular	raised	entire	translucent
3	*Lutiebacter rhizovicinus* AEW 1	variable	rod/cigar	RT	8	yellow	shiny	circular	flat	entire	translucent
3	*Luteibacter sp.* AEZb	mixed	rods/cocci	RT	5	cream	shiny	circular	flat	entire	translucent
3	*Luteibacter rhizovicinus* AFBa	mixed	cigars/rods	RT	5	cream	shiny	circular	flat	filamentous	translucent
3	*Luteibacter rhizovicinus* AFBb	mixed	rods	RT	5	cream	shiny	circular	flat	entire	translucent
4	*Chryseobacterium*	negative	rod/cigar	RT	3	orange	shiny	circular	convex	entire	opaque

	soldanellicola ABH										
4	*Variovorax paradoxus* ABI	negative	rod/cigar	RT	3	yellow	shiny	circular	raised	entire	translucent
4	*Chryseobacterium soldanellicola* ABX	negative	rod	RT	3	orange	shiny	circular	pulvinate	erose	opaque
4	*Chryseobacterium jejuense* ACC1	negative/ variable	rod	RT	3	orange	shiny	punctiform	convex	entire	opaque
4	*Chryseobacterium sp.* ACC2	negative	rod	RT	3	orange	shiny	punctiform	convex	entire	opaque
4	*Chryseobacterium soldanellicola* ACI	negative	rod	RT	3	orange	shiny	punctiform	raised	entire	opaque
4	*Chryseobacterium sp.* ACS	negative	rod/cigar	RT	3	orange	shiny	punctiform	raised	entire	opaque
4	*Chryseobacterium sp.* ACT	negative	cocci/cigar	RT	3	orange	shiny	punctiform	raised	entire	opaque
4	*Chryseobacterium sp.* ACY	negative	rod	RT	3	orange	shiny	punctiform	convex	entire	opaque
4	*Chryseobacterium soldanellicola* ADA	positive	rod/cigar	RT	7	orange	shiny	punctiform	convex	entire	opaque
4	*Chryseobacterium sp.* ADD	positive	rod	RT	7	yellow	shiny	punctiform	raised	entire	translucent
4	*Variovorax paradoxus* ADE	negative	cigar	RT	7	cream	shiny	punctiform	raised	erose	opaque
4	*Chryseobacterium sp.* ADI	positive	rod	RT	7	cream	shiny	circular	raised	entire	translucent
4	*Chryseobacterium soldanellicola* ADO	variable	cigar	RT	7	white	shiny	circular	flat	entire	opaque
4	*Streptomyces sp.* AEA	positive	rod	RT	8	ND	ND	ND	ND	ND	ND
4	*Caulobacter sp.* AEJ	positive	rod	RT	8	cream	shiny	punctiform	flat	entire	opaque
4	*Staphylococcus epidermidis/Pseudomonas sp.* AITb	mixed	mixed	60	o/n	cream	shiny	circular	flat	undulate	opaque
4	*Staphylococcus epidermidis* M2	positive	cocci	50^	o/n	cream	shiny	circular	pulvinate	entire	opaque
4	*Staphylococcus hominis* M1	positive	cocci	50^	o/n	cream	shiny	circular	pulvinate	entire	opaque
5	*Pseudomonas sp.* ACH	variable	rod	RT	3	cream	shiny	circular	raised	entire	opaque
5	*Serratia sp.* ACQ	positive	cocci	RT	3	cream	shiny	punctiform	raised	entire	opaque
5	*Serratia sp.* ACV	negative	cocci	RT	3	cream	shiny	punctiform	convex	entire	opaque
5	*Rhizobium sp.* ADG	variable	cigar	RT	7	white	shiny	circular	raised	entire	opaque
5	*Serratia sp.* ADH	positive	cigar	RT	7	orange	shiny	circular	convex	erose	opaque/translucent edge

5	*Serratia sp.* ADU	positive	cigar/rod	RT	7	cream	shiny	punctiform	flat	erose	opaque/translucent
5	*Advenella sp.* AED	positive	cocci	RT	8	cream	shiny	punctiform	raised	entire	opaque
5	*Sphingomonas sp.* AEH1	positive	rod	RT	8	grey white	shiny	punctiform	flat	entire	translucent
5	*Sphingomonas sp.* AEH2	positive	rod	RT	5	yellow	shiny	punctiform	raised	entire	opaque
5	*Advenella kashmirensis* AEN	variable	cocci	RT	8	ND	ND	ND	ND	ND	ND
5	*Advenella kashmirensis* AET1	positive	rod	RT	8	white	shiny	circular	raised	entire	translucent
5	*Sphingomonas yunnanensis* AET2	positive	rod	RT	8	yellow	shiny	punctiform	flat	entire	opaque
5	*Agrobacterium sp./Rhizobium sp.* AEZa	variable	cigar	RT	5	cream	shiny	circular	flat	entire	translucent
6	*Pseudomonas abietaniphila* ABC	variable	cigar	RT	3	cream	shiny	punctiform	raised	entire	translucent
6	*Pseudomonas abietaniphila* ABD	positive	rod	RT	3	ND	ND	ND	ND	ND	ND
6	*Pseudomonas sp.* ABG	positive	rod	RT	3	white	shiny	circular	raised	entire	opaque/translucent edge
6	*Pseudomonas abietaniphila* ABY	negative	rod	RT	3	cream	shiny	circular	raised	entire	opaque
6	*Pseudomonas sp.* ACA	negative/mixed	rod	RT	3	cream	shiny	circular	raised	entire	translucent
6	*Pseudomonas sp.* ACD	variable	rod/cigar	RT	3	cream	shiny	circular	convex	entire	opaque
6	*Pseudomonas sp.* ACK	positive	rod	RT	3	cream	shiny	circular	raised	entire	opaque
6	*Pseudomonas sp.* ACM	positive	rod	RT	3	cream	shiny	circular	raised	entire	opaque
6	*Pseudomonas sp.* ACO	positive	rod/cigar	RT	3	cream	shiny	circular	flat	entire	opaque
6	*Pseudomonas sp.* ACP	variable	rod	RT	3	cream	shiny	circular	convex	entire	opaque
6	*Pseudomonas sp.* ACU	variable	cocci	RT	3	cream	shiny	punctiform	convex	entire	opaque
6	*Pseudomonas sp.* ACZ	negative	rod	RT	3	cream	shiny	circular	pulvinate	entire	opaque
6	*Pseudomonas abietaniphila* AFEa	negative*	cigar	RT	5	cream	shiny	circular	raised	entire	translucent
6	*Pseudomonas abietaniphila* AFF1	variable	cigar	RT	8	white	shiny	punctiform	raised	entire	opaque
6	*Pseudomonas brenneri* AIE2	positive	rod	40	o/n	cream	shiny	circular	raised	undulate	opaque
6	*Pseudomonas sp.* AIS1a	negative	rod	60	o/n	cream	shiny	circular	flat	undulate	opaque
6	*Pseudomonas sp.* AIS2	positive	rod	60	o/n	cream	shiny	circular	flat	undulate	opaque

6	*Pseudomonas sp.* AIWb	negative	rod	70	o/n	cream	shiny	circular	flat	undulate	opaque
6	*Pseudomonas sp.* AJI	positive	rod	40*	o/n	cream	shiny	circular	flat	undulate	opaque
7	*Bacillus subtilis* AHU	positive	unsure	50	o/n	cream	shiny	punctiform	raised	entire	opaque
7	*Brevibacillus agri* AHW	positive	rod	50	o/n	cream	shiny	circular	convex	entire	opaque
7	*Bacillus licheniformis* AHY1	positive	rod	50	o/n	cream	shiny	irregular	convex	undulate	opaque
7	*Bacillus licheniformis* AHY2	positive	rod	50	o/n	cream	shiny	circular	convex	entire	opaque
7	*Bacillus licheniformis* AIB	variable	rod	50	o/n	cream	shiny	irregular	convex	entire	opaque
7	*Bacillus licheniformis* AID	positive	rod	40	o/n	cream	shiny	circular	raised	undulate	opaque
7	*Bacillus licheniformis* AIE1	positive	rod	40	o/n	cream	dry	circular	convex	filamentous	opaque
7	*Bacillus stratosphericus* AIF	positive	rod	40	o/n	cream	shiny	irregular	raised	undulate	opaque
7	*Bacillus licheniformis* AIH 2	positive	rod	40	o/n	cream	shiny	circular	convex	entire	opaque
7	*Bacillus licheniformis* AIK	positive	rod	40	o/n	cream	shiny	circular	convex	entire	opaque
7	*Bacillus subtilis* AIL	positive	rod	50	o/n	cream	shiny	circular	raised	entire	opaque
7	*Bacillus aestuarii* AITa	negative	rod	60	o/n	cream	dry	irregular	raised	undulate	opaque
7	*Bacillus licheniformis* AJF1	positive	rod	40*	o/n	cream	dry	irregular	flat	undulate	opaque
7	*Bacillus subtilis* AJF2	positive	rod	40*	o/n	cream	shiny	circular	raised	entire	opaque
7	*Bacillus licheniformis* AJG	positive	rod	40*	o/n	cream	dry	irregular	flat	undulate	translucent
7	*Bacillus subtilis* AJH1	positive	rod	40*	o/n	cream	shiny	filamentous	flat	erose	translucent
7	*Bacillus licheniformis* AJH2	positive	rod	40*	o/n	cream	shiny	circular	convex	entire	opaque
7	*Bacillus licheniformis* AJK	positive	rod	40*	o/n	cream	shiny	circular	flat	undulate	translucent
7	*Bacillus subtilis* AJN1a	negative	cigar	40	o/n	cream	shiny	circular	raised	entire	opaque
7	*Bacillus licheniformis* AJN1b	positive	rods	40	o/n	cream	shiny	punctiform	raised	lobate	opaque
7	*Bacillus licheniformis* AJO	positive	rod	50*	o/n	cream	shiny	circular	convex	entire	opaque
7	*Bacillus licheniformis* AJT	positive	rod	50*	o/n	cream	dry	irregular	flat	erose	translucent
7	*Bacillus licheniformis* AJVa	negative	rod	50	o/n	cream	shiny	punctiform	raised	entire	opaque
7	*Bacillus licheniformis* AJVb	negative	rod	50	o/n	cream	shiny	circular	convex	entire	opaque
7	*Brevibacillus agri* AJY	positive	rod	60*	o/n	cream	shiny	circular	convex	entire	opaque
7	*Bacillus subtilis* M3	positive	rod	50	o/n	cream	shiny	circular	raised	entire	opaque
7	*Bacillus coagulans* ABP	positive	rod	40*	o/n	cream	shiny	circular	convex	entire	opaque
7	*Bacillus coagulans* ABQ	positive	rod	40*	o/n	cream	shiny	circular	convex	entire	opaque
7	*Bacillus coagulans* ABR	positive	rod	40*	o/n	cream	shiny	circular	convex	entire	opaque
	ND ABEa	mixed	cigars/rods	RT	5	cream	dry	punctiform	flat	lobate	opaque

ND ACW	negative	cocci	RT	3	cream	shiny	punctiform	raised	entire	opaque
ND ACX	negative	cocci	RT	3	cream	shiny	punctiform	raised	entire	opaque
ND AFF 2	mixed	rods/cigars	RT	8	yellow	shiny	punctiform	flat	entire	translucent
ND AIH 1	positive	rod	40	o/n	cream	shiny	circular	raised	entire	opaque
ND AIS1b	negative	rod	60	o/n	cream	shiny	circular	flat	entire	opaque
ND AIU	positive	rod	60	o/n	cream	shiny	punctiform	raised	undulate	opaque
ND AIWa	negative	rod	70	o/n	cream	shiny	circular	flat	entire	opaque
ND AJP	positive	cocci	50*	o/n	cream	shiny	circular	flat	entire	opaque
ND AKE	positive	cocci	60*	o/n	cream	shiny	circular	flat	undulate	opaque

8.1.3 BLAST statistics for bacterial isolates in Isolation and Identification of Bacterial Library

Table 8.3 Identification of bacteria using the 16S rRNA hypervariable regions via a range of primers.

Two databases were used to identify each bacterium: NCBI BLAST and RDP BLAST. ND – not determined; ID - identity. Unique common oligomers approximately reflects sequence length counting only unique oligos to compensate for bias (Center for Microbial Ecology, 2012).

			NCBI BLAST			RDP BLAST	
Isolate code	**16S ID**	**Query length (bp)**	**Query coverage %**	**ID %**	**E Value**	**Seqmatch score (S_ab) %**	**Unique Common Oligomers**
ABC	*Pseudomonas abietaniphila*	1222	100	99	0.0	99.5	1407
ABD	*Pseudomonas abietaniphila*	1210	100	99	0.0	99.5	1407
ABEa	*ND*	ND	ND	ND	ND	ND	ND
ABEb	*Burkholdeia fungorum*	1349	100	100	0.0	100	1385
ABF 1a	*Burkholdeia fungorum*	1337	100	100	0.0	100	1385
ABF 1b	*Burkholdeia fungorum*	1343	100	100	0.0	100	1385
ABF 2	*Enhydrobacter sp./Moraxella osloensis*	1251	100/100	99/99	0.0/0.0	100/100	1243/1433
ABG	*Pseudomonas sp.*	1238	98	98	0.0	97.1	0664
ABH	*Chryseobacterium soldanellicola*	912	100	98	0.0	98.3	1280
ABI	*Variovorax paradoxus*	1217	100	99	0.0	99.4	1431
ABP	*Bacillus coagulans*	955	99	99	0.0	100	0688
ABQ	*Bacillus coagulans*	880	98	99	0.0	99.4	0688
ABR	*Bacillus coagulans*	600	98	99	0.0	100	0688
ABX	*Chryseobacterium soldanellicola*	1195	100	98	0.0	99	1280
ABY	*Pseudomonas abietaniphila*	1216	100	100	0.0	100	1357
ACA	*Pseudomonas sp.*	1198	100	98	0.0	96.6	1357
ACB	*Enhydrobacter aerosaccus/Moraxella sp.*	1224	100/100	99/99	0.0/0.0	100/100	0542/1240
ACC 1	*Chryseobacterium jejuense*	1201	100	99	0.0	99.8	1365
ACC 2	*Chryseobacterium sp.*	1194	100	99	0.0	99.3	1369
ACD	*Pseudomonas sp.*	1214	100	99	0.0	100	1404
ACH	*Pseudomonas sp.*	236	47	86	2e-29	30.9	1470
ACI	*Chryseobacterium soldanellicola*	1322	99	99	0.0	99.1	1280
ACK	*Pseudomonas sp.*	1218	100	98	0.0	97.1	0664
ACL	*Pedobacter terrae*	1209	100	99	0.0	95.7	1383
ACM	*Pseudomonas sp.*	1221	100	98	0.0	96.7	1357
ACN	*Luteibacter sp.*	1362	99	98	0.0	99.8	0901
ACO	*Pseudomonas sp.*	233	93	77	5e-37	28.8	1342

ACP	*Pseudomonas sp.*	1221	100	99	0.0	100	1404
ACQ	*Serratia sp.*	400	94	96	2e-59	72.6	1443
ACS	*Chryseobacterium sp.*	1204	100	99	0.0	99.3	1369
ACT	*Chryseobacterium sp.*	1196	100	99	0.0	99.3	1369
ACU	*Pseudomonas sp.*	280	81	91	6e-87	54.6	1403
ACV	*Serratia sp.*	440	89	97	1e-134	76.5	1444
ACW	*ND*	ND	ND	ND	ND	ND	ND
ACX	*ND*	ND	ND	ND	ND	ND	ND
ACY	*Chryseobacterium sp.*	1209	100	99	0.0	99.3	1369
ACZ	*Pseudomonas sp.*	1234	100	99	0.0	100	1404
ADA	*Chryseobacterium soldanellicola*	1194	100	99	0.0	95.2	1368
ADB	*Luteibacter rhizovicinus*	1359	100	99	0.0	99.6	1387
ADC	*Burkholderia sordidicola*	1195	100	99	0.0	98.8	1283
ADD	*Chryseobacterium sp.*	1211	100	99	0.0	99.8	1365
ADE	*Variovorax paradoxus*	1192	100	99	0.0	99.4	1431
ADF	*Lutiebacter rhizovicinus*	1200	100	100	0.0	100	1387
ADG	*Rhizobium sp.*	1288	100	99	0.0	99.2	1328
ADH	*Serratia sp.*	1225	100	99	0.0	99.8	0556
ADI	*Chryseobacterium sp.*	1218	100	99	0.0	99.3	1369
ADJa	*Burkholderia sp.*	1203	100	98	0.0	100	1385
ADJb	*Burkholderia sp.*	1189	100	99	0.0	100	1366
ADK	*Enhydrobacter sp./Moraxella sp.*	1217	100/100	99/99	0.0/0.0	100/100	1433/1490
ADM	*Micrococcus luteus*	1174	99	99	0.0	100	0560
ADN	*Luteibacter rhizovicinus*	1354	100	100	0.0	100	1387
ADO	*Chryseobacterium soldanellicola*	1200	100	99	0.0	99	1280
ADP	*Luteibacter sp.*	200	98	97	1e-89	91.1	1423
ADQ	*Microbacterium sp.*	1329	99	99	0.0	98.8	1418
ADR	*Burkholderia fungorum*	1214	100	100	0.0	100	1385
ADS	*Microbacterium sp.*	1326	100	99	0.0	99.0	1418
ADT	*Phyllobacterium ifriqiyense*	1156	100	100	0.0	100	1338
ADU	*Serratia sp.*	624	80	97	1e-77	73.3	1369
ADV	*Burkholderia sordidicola*	1216	100	99	0.0	97.7	1415
ADW	*Massilia dura*	1350	100	99	0.0	92.9	1410
ADX 1	*Burkholderia sordidicola*	1214	100	99	0.0	97.7	1415
ADX 2	*Luteibacter rhizovicinus*	1208	100	100	0.0	96.1	0735
AEA	*Streptomyces sp.*	681	100	100	0.0	100	1410
AEB (M4)	*Luteibacter rhizovicinus*	1227	100	99	0.0	99.3	1365
AEC	*Luteibacter sp.*	1360	100	97	0.0	94.6	0735
AED	*Advenella sp.*	1226	100	99	0.0	100	1360
AEH 1	*Sphingomonas sp.*	1174	100	97	0.0	94.4	0429
AEH 2	*Sphingomonas sp.*	1170	100	97	0.0	96	0429
AEIa	*Burkholdeia fungorum*	1349	100	100	0.0	100	1385
AEIb	*Burkholdeia*	1328	100	99	0.0	99.7	1385

	fungorum						
AEJ	*Caulobacter sp.*	1288	100	99	0.0	98.3	1347
AEK	*Luteibacter sp.*	1163	100	99	0.0	96	1337
AELa	*Burkholderia fungorum*	1227	99	100	0.0	99.6	1385
AELb	*Burkholderia fungorum*	1217	100	100	0.0	100	1385
AEM	*Burkholderia sp.*	1339	100	97	0.0	94.0	0923
AEN	*Advenella kashmirensis*	1251	98	99	0.0	98.1	1339
AEOa	*Luteibacter rhizovicinus*	1389	100	100	0.0	100	1387
AEOb	*Luteibacter rhizovicinus*	1389	100	100	0.0	100	1387
AEP	*Microbacterium sp.*	1283	92	96	0.0	97.7	0393
AEQ	*Herbaspirillum hiltneri*	1347	100	99	0.0	97.8	1392
AER	*Luteibacter rhizovicinus*	1200	100	100	0.0	100	1387
AES 1	*Burkholderia fungorum*	1224	100	100	0.0	100	1385
AES 2	*Luteibacter rhizovicinus*	1255	99	95	0.0	95.9	0735
AET 1	*Advenella kashmirensis*	1217	100	99	0.0	100	1339
AET 2	*Sphingomonas yunnanensis*	1319	100	99	0.0	95.6	1338
AEU	*Microbacterium sp.*	1335	100	99	0.0	98.6	1418
AEV	*Luteibacter rhizovicinus*	1257	99	96	0.0	95.0	0735
AEW 1	*Lutiebacter rhizovicinus*	1220	100	97	0.0	99.4	1387
AEW2a	*Burkholdeia fungorum*	1357	100	100	0.0	100	1385
AEW2b	*Burkholdeia fungorum*	1201	100	100	0.0	100	1385
AEXa	*Burkholderia fungorum*	1219	100	100	0.0	100	1385
AEXb	*Burkholderia fungorum*	1215	100	100	0.0	100	1385
AEYa	*Burkholdeia fungorum*	1203	100	100	0.0	100	1385
AEYb	*Burkholdeia fungorum*	1189	100	100	0.0	100	1385
AEZa	*Agrobacterium sp./Rhizobium sp.*	1154	100/100	99/99	0.0/0.0	100/99.5	1205/1265
AEZb	*Luteibacter sp.*	1333	94	93	0.0	92.1	0735
AFA	*Burkholderia fungorum*	1225	100	100	0.0	100	1385
AFBa	*Luteibacter rhizovicinus*	1393	100	100	0.0	100	1387
AFBb	*Luteibacter rhizovicinus*	1391	100	100	0.0	100	1387
AFC	*Burkholderia fungorum*	1200	100	100	0.0	100	1385
AFDa	*Burkholderia fungorum*	1163	100	99	0.0	99	1385

AFDb	*Burkholderia fungorum/L. rhizovicinus*	1197	100/97	100/97	0.0/0.0	100/80.9	1385/1384
AFEa	*Pseudomonas abietaniphila*	1200	100	100	0.0	100	1357
AFEb	*Burkholdeia fungorum*	1361	100	100	0.0	100	1385
AFF 1	*Pseudomonas abietaniphila*	1220	100	100	0.0	100	1357
AFF 2	*ND*	ND	ND	ND	ND	ND	ND
AFG	*Dyella sp.*	1368	99	99	0.0	99.5	1433
AHU	*Bacillus subtilis*	1306	100	100	0.0	100	1423
AHW	*Brevibacillus agri*	1346	100	99	0.0	100	1365
AHY1	*Bacillus licheniformis*	1202	100	98	0.0	91.3	1358
AHY2	*Bacillus licheniformis*	1355	100	100	0.0	100	1465
AIB	*Bacillus licheniformis*	1309	100	100	0.0	100	1463
AID	*Bacillus licheniformis*	1362	100	99	0.0	99.2	1122
AIE1	*Bacillus licheniformis*	1350	98	93	0.0	89.3	0541
AIE2	*Pseudomonas brenneri*	1253	93	96	0.0	100	0533
AIF	*Bacillus pumilus*	1220	100	98	0.0	100	0814
AIH1	*ND*	ND	ND	ND	ND	ND	ND
AIH2	*Bacillus licheniformis*	1300	99	97	0.0	100	0692
AIK	*Bacillus licheniformis*	1299	100	99	0.0	100	1427
AIL	*Bacillus subtilis*	1242	100	100	0.0	100	1423
AIS1a	*Pseudomonas sp.*	1302	91	96	0.0	81.6	1126
AIS1b	*ND*	ND	ND	ND	ND	ND	ND
AIS2	*Pseudomonas sp.*	1304	96	95	0.0	100	0580
AITa	*Bacillus aestuarii*	1400	99	97	0.0	99	1385
AITb	*Staphylococcus epidermidis/ Pseudomonas sp.*	1389	100/97	99/95	0.0/0.0	100/100	0434/0533
AIU	*ND*	ND	ND	ND	ND	ND	ND
AIWa	*ND*	ND	ND	ND	ND	ND	ND
AIWb	*Pseudomonas sp.*	1297	97	96	0.0	100	1409
AJF1	*Bacillus licheniformis*	1317	100	100	0.0	100	1463
AJF2	*Bacillus subtilis*	1306	100	99	0.0	100	0691
AJG	*Bacillus licheniformis*	1218	100	99	0.0	99.8	0664
AJH1	*Bacillus subtilis*	1359	100	99	0.0	100	0691
AJH2	*Bacillus licheniformis*	1355	100	100	0.0	100	1465
AJI	*Pseudomonas sp.*	1299	98	96	0.0	100	1232
AJK	*Bacillus licheniformis*	1354	85	92	0.0	80.4	0588
AJN1a	*Bacillus subtilis*	1308	99	96	0.0	99	0597
AJN1b	*Bacillus licheniformis*	1309	100	99	0.0	100	1461
AJO	*Bacillus licheniformis*	1366	100	99	0.0	100	1427
AJP	*ND*	ND	ND	ND	ND	ND	ND
AJT	*Bacillus licheniformis*	1350	99	96	0.0	100	0701
AJVa	*Bacillus licheniformis*	1387	100	99	0.0	100	0791
AJVb	*Bacillus licheniformis*	1382	100	99	0.0	100	1467
AJY	*Brevibacillus agri*	1366	100	99	0.0	99	1415
AKE	*ND*	ND	ND	ND	ND	ND	ND
M1	*Staphylococcus hominis*	1236	100	99	0.0	100	1457
M2	*Staphylococcus epidermidis*	1314	95	97	0.0	100	0784
M3	*Bacillus subtilis*	1235	100	100	0.0	100	1423

8.1.4 Carbon utilisation profiles of mesophiles and thermotolerant isolates

Table 8.4 Mesophilic bacterial isolates carbon utilisation profiles.

Cells are coloured via intensity, the larger the relative activity the darker the colour used. Cellulase (CMC - Carboxymethylcellulose) activity is presented in red; xylanase activity is coloured blue, and starch activity is coloured green. Each assay was conducted at 4 temperatures. Clade relates to phylogenetic tree in Figure 3.6. Relative activity scale: 0 showed no activity halo; 1 represented 1-2 mm halo; 2 was 3-5 mm halo; 3 was 5-7 mm; 4 represented 8-10 mm halo; 5 was 11-13 mm halo, and 6 showed over activity of 14 mm or more (where the halo expanded out of the grid square).

		CMC					Xylan					Starch				
Clade	**16S ID**	**pH 4**	**pH 5.5**	**pH 6.2**	**pH 8**	**pH 10**	**pH 4**	**pH 5.5**	**pH 6.2**	**pH 8.5**	**pH 11**	**pH 4**	**pH 5.5**	**pH 6.2**	**pH 8**	**pH 9**
1	*Burkholdeia fungorum ABEb*	0	2	2	1	0	0	2	1	0.5	0	1	1	1	1	0
1	*Burkholdeia fungorum ABF1a*	0	1.5	1	2	0	0	1.5	1	1	0	1	1	1	0.5	0
1	*Burkholdeia fungorum ABF1b*	0	1.5	1	2	0	0	1.5	1	0.5	0	1	1	1	1	0
1	*Burkholderia sordidicola ADC*	0	3	1	0	0	0	2	1	1.5	0	0	0.5	1	0	0
1	*Burkholderia sp. ADJa*	0	3	3	0.5	0	0	3	1	1	0	1	2	1	0.5	0
1	*Burkholderia sp. ADJb*	0	2.5	2	1	0	0	3	1	1	0	1	1.5	1	0.5	0
1	*Burkholderia fungorum ADR*	0	1	0	0	0	0	1	0	0	0	0	0.5	0	0	0
1	*Burkholdeia fungorum AEIa*	0	1.5	1	1	0	0	2	1	1	0	1	1	1	0.5	0.5
1	*Burkholdeia fungorum AEIb*	0	1.5	1	1	0	0.5	1.5	1	1	0	1	1	1	1	0
1	*Burkholderia fungorum AELa*	0	1.5	1	1	0	0.5	1.5	1	1	0	1.5	1	1	1	0
1	*Burkholderia fungorum AELb*	0	2	1	1	0	0	1.5	1	1	0	1	1	0.5	1	0
1	*Burkholderia sp. AEM*	0	0.5	1	1	0	0	1	1	0	0	1	0.5	1	0	0
1	*Burkholderia sordidicola ADV*	0	1	0	0	0	0	1	0	0	0	0	0.5	1	0	0
1	*Burkholderia sordidicola ADX 1*	0	1	0	0	0	0	1.5	1	1	0	0	1	1	0	0
1	*Burkholderia fungorum AES 1*	0	1.5	1	0	0	0	2	1	0	0	0	1	1	0	0
1	*Burkholdeia fungorum AEW2a*	0	1.5	1	0	0	1	1.5	1	1	0	1	1	1	1	0
1	*Burkholdeia fungorum AEW2b*	0	1	1	0	0	0	1.5	1	1	0	1	1	1	1	0

1	*Burkholderia fungorum AEXa*	0	2	1	0	0	0	2	1	1	0	1	1	1	1	0
1	*Burkholderia fungorum AEXb*	0	3	1	1	0	0	2	1	1	0	1	1	1	1	0
1	*Burkholdeia fungorum AEYa*	0	1.5	1	0.5	0	0	2.5	1	1	0	1.5	1	1	1	0
1	*Burkholdeia fungorum AEYb*	0	1.5	1	0.5	0	0	2	1	1	0	1	1	1	0.5	0
1	*Burkholderia fungorum AFA*	0	1	0	0	0	0	1	0	0	0	0	0.5	0	0	0
1	*Burkholderia fungorum AFC*	0	1	0	0	0	0	1.5	0	0	0	0	1	0.5	0.5	0
1	*Burkholderia fungorum AFDa*	0	0	1	1	0	1	0.5	1	0.5	0	1	0.5	1	1	0
1	*Burkholderia fungorum/L. rhizovicinus AFDb*	0	1	1	1.5	0	0	1.5	1	0.5	0	1	1	1	1	0
1	*Burkholdeia fungorum AFEb*	0	4	1	0	0	0	3	1	1.5	0	1	1	1	1	0
2	*Enhydrobacter sp. Or Moraxella osloensis ABF2*	0	1.5	1	0	0	0	0.5	1	1	0	0	0.5	0	0	0
2	*E. aerosaccus/Moraxella sp. ACB*	0	3	0	0	0	0	1.5	0	0	0	0	1	0	0	0
2	*Pedobacter terrae ACL*	0	4	3	1	0	0	2.5	3	3	0	0	2.5	2	2	0
2	*Enhydrobacter sp./Moraxella sp. ADK*	0	1.5	0	0	0	0	1.5	1	1	0	0	0.5	0.5	0	0
2	*Micrococcus luteus ADM*	0	5	3	0	0	0	5	3	1	0	0	1.5	2	0	0.5
2	*Microbacterium sp. ADQ*	0	0.5	0	0	0	0	1	0.5	0	0	0	0.5	0.5	0	0
2	*Microbacterium sp. ADS*	0	1.5	0	0	0	0	2	0	0	0	0	0.5	1	0	0
2	*Phyllobacterium ifriqiyense ADT*	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0
2	*Massilia dura ADW*	0	0.5	0	0	0	0	0.5	1	0	0	0	0.5	0.5	0	0
2	*Microbacterium sp. AEP*	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0.5
2	*Herbaspirillum hiltneri AEQ*	0	0.5	0	0	0	0	1	0.5	0	0	0	0	0.5	0	0
2	*Microbacterium sp. AEU*	0	0	0	0	0	0	0	0.5	0	0	0	0	0.5	0	0
2	*Dyella sp. AFG*	0	3.5	2.5	0	0	0	3	4	0	0	0	0.5	0.5	0	0
3	*Luteibacter sp. ACN*	0	4.5	0	0	0	0	3.5	0	0	0	0	1	0	0	0
3	*Luteibacter rhizovicinus ADB*	0	5	0	0	0	0	4.5	0	0	0	0	2.5	0	0	0
3	*Lutiebacter rhizovicinus ADF*	0	0	0	0	0	0	0.5	0	0	0	0	0	0.5	0	0
3	*Luteibacter rhizovicinus ADN*	0	0.5	0	0	0	0	0.5	0	0	0	0	0	0	0	0
3	*Luteibacter sp. ADP*	0	2.5	0	0	0	0	3	0	0	0	0	2	0	0	0
3	*Luteibacter rhizovicinus ADX 2*	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0	0

3	*Luteibacter rhizovicinus AEB (M4)*	0	2.5	1	0	0	0	2	0	0	0	0	0	0.5	0	0
3	*Luteibacter sp. AEC*	0	0.5	0	0	0	0	0.5	0.5	0	0	0	0	0.5	0	0
3	*Luteibacter sp. AEK*	0	0	0	0	0	0	0	0.5	0	0	0	0	0.5	0	0
3	*Luteibacter rhizovicinus AEOa*	0	1.5	1	1.5	0	0	2	1	1	0	1	1	1	1	0
3	*Luteibacter rhizovicinus AEOb*	0	1.5	1	1.5	0	1.5	2	1	1	0	1	2	1	1	0.5
3	*Luteibacter rhizovicinus AER*	0	0.5	0	0	0	0	1	0	0	0	0	0	0.5	0	0
3	*Luteibacter rhizovicinus AES 2*	0	0.5	1	0	0	0	1	0.5	0	0	0	0	1	0	0.5
3	*Luteibacter rhizovicinus AEV*	0	0.5	0	0	0	0	1.5	1	0	0	0	0	1	0	0
3	*Lutiebacter rhizovicinus AEW 1*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	*Luteibacter sp. AEZb*	0	1.5	1	0	0	0	1.5	1	1	0	1	1	1	0	0
3	*Luteibacter rhizovicinus AFBa*	0	2	0.5	0	0	0	2	0.5	1	0	1	1	0.5	0.5	0
3	*Luteibacter rhizovicinus AFBb*	0	2	1	0	0	0	2.5	1	1	0	1	1	0	1	0
4	*Chryseobacterium soldanellicola ABH*	0	5	5	0	0	0	3.5	5	5.5	0	0	4.5	3	4	0
4	*Variovorax paradoxus ABI*	0	0.5	1	0	0	0	1	1	1	0	0	1	1	0	0
4	*Chryseobacterium soldanellicola ABX*	0	5	5	0	0	0	4	5	5	0	0	4	4	4	0
4	*Chryseobacterium jejuense ACC1*	0	5	5	0	0	0	4	5	6	0	0	4.5	3	4.5	0
4	*Chryseobacterium sp. ACC2*	0	5	5	1	0	0	4	5	6	0	0	3	3	4.5	0
4	*Chryseobacterium soldanellicola ACI*	0	5	4.5	0	0	0	4	5	5	0	0	3.5	3	3	0
4	*Chryseobacterium sp. ACS*	0	5	5	0	0	0	4	5	5	0	0	4	3	4.5	1
4	*Chryseobacterium sp. ACT*	0	5	5	0	0	0	3.5	5	3.5	0	0	3.5	3.5	5	0
4	*Chryseobacterium sp. ACY*	0	5	5	0	0	0	4.5	5	5	0	0	3.5	3	4.5	1
4	*Chryseobacterium soldanellicola ADA*	0	4	3.5	0	0	0	3	3	1.5	0	0	1	2	1.5	0
4	*Chryseobacterium sp. ADD*	0	4	5	0	0	0	3	5	5	0	0	2.5	3	4.5	0
4	*Variovorax paradoxus ADE*	0	0.5	0	0.5	0	0	1	1	5	0	0	1	1	3	0
4	*Chryseobacterium sp. ADI*	0	3	5	0	0	0	3	5	5	0	0	4	3	5	0
4	*Chryseobacterium soldanellicola ADO*	0	3	1	0	0	0	3	1	1	0	0	2.5	1	0	0
4	*Streptomyces sp. AEA*	0	3.5	2	2	0	0	2.5	2	2.5	0	0	2	2	2	0
4	*Caulobacter sp. AEJ*	0	0.5	0	0	0	0	1	0	0	0	0	1	0.5	0	0

4	*Staphylococcus epidermidis M2*	0	2.5	0	0	0	0	2.5	0	0	0	0	0	0	0	0
4	*Staphylococcus hominis M1*	0	5	5	0	0	0	5	5	5.5	0	0	4	4	4	0
5	*Pseudomonas sp. ACH*	0	4	0	0.5	0	0	4	1	3	0	0	1.5	0	0.5	0.5
5	*Serratia sp. ACQ*	0	2	0	0	0	0	1.5	1	1	0	0	0	0.5	0	1
5	*Serratia sp. ACV*	0	2.5	0	0	0	0	2.5	1	1	0	0	0.5	0	0	0.5
5	*Rhizobium sp. ADG*	0	0.5	0	0	0	0	0.5	0	0	0	0	0.5	0	0	0
5	*Serratia sp. ADH*	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0.5
5	*Serratia sp. ADU*	0	0.5	0	0	0	0	1	1	1	0	0	0.5	1	0	0
5	*Advenella sp. AED*	0	1	1	1.5	0	0	1	1	1	0	0	0.5	1	1	0.5
5	*Sphingomonas sp. AEH1*	0	0	0	0	0	0.5	0	0	0	0	0	0	0.5	0	0
5	*Sphingomonas sp. AEH2*	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
5	*Advenella kashmirensis AEN*	0	0.5	1	1	0	0	1	1	1	0	0	0.5	1	1	0
5	*Advenella kashmirensis AET1*	0	1	0	0	0	0	1	1	1	0	0	0.5	1	1	0
5	*Sphingomonas yunnanensis AET2*	0	0	0	0	0	0	0.5	0	0	0	0	0	0.5	0	0
5	*Agrobacterium sp./Rhizobium sp. AEZa*	0	1.5	1	0	0	1.5	2	1	1	0	1	1	0.5	0.5	0
6	*Pseudomonas abietaniphila ABC*	0	4.5	3	1	0	0	4	3	4.5	0	0	2	2	2	0
6	*Pseudomonas abietaniphila ABD*	0	4.5	0	0	0	0	4	0	0	0	0	1	0	0	0
6	*Pseudomonas sp. ABG*	0	0.5	0	0	0	0	1	1	4	0	0	0	1	0	0
6	*Pseudomonas abietaniphila ABY*	0	4	2	0	0	0	4	2	4	0	0	1.5	1	1.5	0
6	*Pseudomonas sp. ACA*	0	0.5	1	0	0	0	1.5	1	3	0	0	0.5	1	0	0
6	*Pseudomonas sp. ACD*	0	2.5	1	0	0	0	2	1	2.5	0	0	1	0.5	0	1
6	*Pseudomonas sp. ACK*	0	1.5	1	0	0	0	1	1	3.5	0	0	0	0	0	0
6	*Pseudomonas sp. ACM*	0	3	1	0	0	0	2.5	1	1	0	0	0	1	0	0
6	*Pseudomonas sp. ACO*	0	1.5	0	0	0	0	1	1	1	0	0	0.5	0.5	0	0
6	*Pseudomonas sp. ACP*	0	2.5	0	0.5	0	0	2	1	2	0	0	0.5	1	0	0.5
6	*Pseudomonas sp. ACU*	0	5	0	0	0	0	4.5	1	1	0	0	2	0	0	0
6	*Pseudomonas sp. ACZ*	0	2	0	1	0	0	1.5	1	2.5	0	0	0.5	0.5	0	0.5
6	*Pseudomonas abietaniphila AFEa*	0	5	2	0	0	0	5	1.5	4	0	1	2	1	2	0

6	*Pseudomonas abietaniphila AFF1*	0	3	1	0	0	0	3	1	0	0	0.5	1	1	0	0
7	*Bacillus subtilis M3*	0	4.5	0	0	0	0	5	0	0	0	0	0	0	0	0
7	*Bacillus coagulans ABP*	0	2.5	0	0	0	0	0	0	0	0	0	5	5	1.5	0.5
7	*Bacillus coagulans ABQ*	0	2	0	0	0	0	3	0	0	0	0	5	5	2	1.5
7	*Bacillus coagulans ABR*	0	2.5	0	0	0	0	0	0	0	0	0	5	5	1.5	0
	Unknown sp. ABEa	0	5	5	0	0	0	5	5	5	0	0	5	2	5	0
	Unknown sp. ACW	0	2.5	0	0	0	0	2.5	1	1.5	0	0	0.5	0	0.5	0
	Unknown sp. ACX	0	2	0	0	0	0	1	1	1	0	0	0.5	0	0	0
	Unknown sp. AFF 2	0	2	1	0	0	0	2.5	0.5	0	0	0	1	0.5	0	0

Table 8.5 Thermotolerant bacterial isolates carbon utilisation profiles.

Cells are coloured via intensity, the larger the relative activity the darker the colour used. Cellulase (CMC - Carboxymethylcellulose) activity is presented in red; xylanase activity is coloured blue, and starch activity is coloured green. Each assay was conducted at 4 temperatures. Clade relates to phylogenetic tree in Figure 3.6. Relative activity scale: 0 showed no activity halo; 1 represented 1-2 mm halo; 2 was 3-5 mm halo; 3 was 5-7 mm; 4 represented 8-10 mm halo; 5 was 11-13 mm halo, and 6 showed over activity of 14 mm or more (where the halo expanded out of the grid square).

Clade	16S rRNA identification	CMC				Xylan				Starch				
		pH 4	pH 6.2	pH 8	pH 10	pH 4	pH 6.2	pH 8.5	pH 11	pH 4	pH 6.2	pH 8	pH 9	Temperature (°C)
4	*S. epidermidis/Pseudomonas sp. AITb*	0	0	0	0	0	0	0	0	0	0	0	0	40
		0	0	0	0	0	0	0	0	0	0	0	0	50
		0	0	0	0	0	0	0	0	0	0	0	0	60
		0	0	0	0	0	0	0.5	0	0	0	0	0	70
6	*Pseudomonas brenneri AIE2*	0	0	0	0	0	1	1	0	0	0	0	0	40
		0	0	0	0	0	1	0	0	0	0	3	0	50
		0	0	0	0	0	0	0	0	0	0	0	0	60
		0	0	0	0	0	0	0	0	0	0	0	0	70
6	*Pseudomonas sp. AIS1a*	0	0	0	0.5	0	1	0	0	0	0	0	0	40
		0	1	0	0	0	2	0	0	0	0	0	0	50
		0	2	1.5	0	0	1.5	1	0	0	1	0	0	60
		0	1.5	3	0	0	1.5	0.5	0	0	1	1.5	0	70
6	*Pseudomonas sp. AIS2*	0	0	0	0	0	0.5	0	0	0	0	0	0	40
		0	0	0.5	0	0	0.5	0	0	0	0	0	0	50
		0	0	0	0	0	0	0	0	0	0	0	0	60
		0	0	0	0	0	0	0	0	0	0	0	0	70
6	*Pseudomonas sp. AIWb*	0	0	0	0	0	0	0	0	0	0	0	0	40
		0	1	0	0	0	1	0	0	0	0	0	0	50

		0	1	0	0	0	1	0	0	0	0	0	0	60
		0	1	0	0	0	1	0	0	0	0	0	0	70
6	*Pseudomonas sp. AJI*	0	5	0	0.5	0	5	0	0	0	3	0	0	40
		0	5	0	0	0	5	0	0	0	3	1.5	0	50
		0	1.5	0	0	0	1	0	0	0	0	0	0	60
		0	1	0	0	0	1.5	0	0	0	0	0	0	70
7	*Bacillus subtilis AHU*	0	5.5	4.5	1	0	5.5	5.5	0.5	0	5	4.5	3.5	40
		0	5	5	0	0	0	5.5	1	0	4	4.5	3	50
		0	0.5	2.5	0	0	0	2.5	0.5	0	0	0.5	0	60
		0	0.5	2	0	0	1	1	0	0	0	1.5	0	70
7	*Brevibacillus agri AHW*	0	0	1	0	0	0	0	0	0	0	0	0	40
		0	0	1	0	0	0	0	0.5	0	0	0	0	50
		0	0.5	0	0	0	0	0	0	0	0	0	0	60
		0	0	0	0	0	0	0	0	0	0	0	0	70
7	*Bacillus licheniformis AHY1*	0	4	5	0	0	4	5	0	0	1	3.5	4	40
		0	3	5	0	0	5	5	5	0	0	5	4.5	50
		0	1	3.5	0	0	0.5	5	0	0	0	2	0	60
		0	2	0	0	0	0.5	0.5	0	0	0	0	0	70
7	*Bacillus licheniformis AHY2*	0	4	5	0	0	6	5	0	0	1	5.5	2.5	40
		0	3	5	0	0	0	5	4	0	0	5	4.5	50
		0	1	1.5	0	0	0	4.5	0.5	0	0	0.5	0	60
		0	2.5	0	0	0	1	0	0	0	0	1	0	70
7	*Bacillus licheniformis AIB*	0	4.5	3	0.5	0	5	5	0	0	3	2	1.5	40
		0	5	4.5	0	0	0	4.5	3	0	3.5	4	2	50
		0	0	1.5	1	0	0	2.5	0	0	0	0.5	0	60
		0	1	2.5	0	0	0.5	0.5	0	0	0	1.5	0	70
7	*Bacillus licheniformis AID*	0	3.5	5	0	0	3	4	0	0	0.5	4	2.5	40
		0	2	5	0	0	3	5	4	0	0	5	3	50

		0	1	2	0	0	1	4	0	0	0	1.5	0	60
		0	2	0	0	0	0	0	0	0	0	0	0	70
7	*Bacillus licheniformis AIE1*	0	5	5	0	0	4.5	5	0	0	2	5.5	3.5	40
		0	4	5	0	0	4	5	3	0	1	5	4	50
		0	1	1.5	0	0	2	5	0	0	0	2.5	0	60
		0	1	0	0	0	1.5	0	0	0	0	0	0	70
7	*Bacillus stratosphericus AIF*	0	0	0	0	0	1	1	1	0	0	0	0	40
		0	0	0	0	0	1	1	0	0	0	0	0	50
		0	0	0	0	0	0	0	0	0	0	0	0	60
		0	0	0	0	0	0	0	0	0	0	0	0	70
7	*Bacillus licheniformis AIH 2*	0	5	5	0	0	4.5	5	0	0	1	4	3.5	40
		0	3	5	0	0	4	5	0	0	0.5	6	3.5	50
		0	2	0.5	0	0	0	4.5	0	0	0	1	0	60
		0	2	2	0	0	2.5	0	0	0	0	0	0	70
7	*Bacillus licheniformis AIK*	0	0	0	0	0	1	1	2	0	0	0	0	40
		0	0.5	0	0	0	1	1	0	0	0	0	0	50
		0	0	0	0	0	0	0	0	0	0	0	0	60
		0	0	0	0	0	0	0	0	0	0	0	0	70
7	*Bacillus subtilis AIL*	0	4	4.5	0.5	0	3	4	0	0	2	3	2	40
		0	4	5	0	0	4	4	0	0	2	3	4	50
		0	1	0.5	0	0	2	3	0.5	0	0	1	0	60
		0	1.5	2	0	0	2	0	0	0	0	0.5	0	70
7	*Bacillus aestuarii AITa*	0	0	0	0	0	0	0	0	0	0	0	0	40
		0	0	0	0	0	0	0	0	0	0	0	0	50
		0	0	0	0	0	0	0	0	0	0	0	0	60
		0	0	0	0	0	0	0	0	0	0	0	0	70
7	*Bacillus licheniformis AJF1*	0	4	5	0	0	4.5	5	0	0	2	5.5	4.5	40
		0	2	5	0	0	3.5	5	2.5	0	1	5.5	5	50

		0	1	1.5	0	0	0.5	5	0	0	0	1.5	0.5	60
		0	1	0	0	0	1	0	0	0	0	0	0	70
7	*Bacillus subtilis AJF2*	0	5	5	3	0	5	5.5	1.5	0	3	4	4	40
		0	5	5	1	0	0	5	2.5	0	5.5	4	2.5	50
		0	0.5	2	0	0	1.5	0	1	0	0	0	0	60
		0	1	2	0	0	1	0	0	0	0	1	0	70
7	*Bacillus licheniformis AJG*	0	4	5	1	0	3	5	2	0	0.5	5.5	5.5	40
		0	3	5	0	0	0	5.5	2.5	0	0	5.5	5	50
		0	1	1.5	0	0	1	1.5	0.5	0	0	0	0	60
		0	2.5	1	0	0	1	0	0	0	0	1	0	70
7	*Bacillus subtilis AJH1*	0	5	5	1	0	4	5	1.5	0	2	5.5	6	40
		0	4	5	2	0	0	5	1.5	0	1	6	6	50
		0	0	3.5	0	0	1	5	0	0	0	2.5	0	60
		0	1	0.5	0	0	1.5	2	0	0	0	0	0	70
7	*Bacillus licheniformis AJH2*	0	5.5	3.5	0	0	5	4	0	0	3.5	4	2.5	40
		0	5	4	0	0	5	4	0.5	0	3	4.5	2	50
		0	0	1	0	0	1	0	0	0	0	0	0	60
		0	0.5	2	0	0	1	0	0	0	0	0	0	70
7	*Bacillus licheniformis AJK*	0	4	5	0.5	0	4	5	0	0	1	2.5	3.5	40
		0	4	5	0	0	0	5	4.5	0	0	4	4	50
		0	1	4.5	0	0	0.5	5	0.5	0	0	2	0	60
		0	1	0	0	0	0.5	0	0	0	0	1.5	0	70
7	*Bacillus subtilis AJN1a*	0	5	5	1	0	5.5	5	2	0	4	3.5	4	40
		0	5	5	0	0	0	5	1.5	0	5	4	3	50
		0	1	2.5	0.5	0	1.5	3.5	0	0	0	1	0.5	60
		0	1	2	0	0	1.5	1	0	0	0	1.5	0	70
7	*Bacillus licheniformis AJN1b*	0	4	4.5	1	0	5	4	0.5	0	3.5	3.5	3.5	40
		0	4.5	5	1	0	0	4.5	0.5	0	2	4	2.5	50

		0	0.5	2.5	0	0	1.5	3	0	0	0	0.5	0	60
		0	0.5	2	0	0	1	2	0	0	0	1.5	0	70
7	*Bacillus licheniformis AJO*	0	5	5	0.5	0	4	4.5	0	0	0.5	4	3	40
		0	4	5	0	0	4	5.5	0	0	0	5	5	50
		0	0	3	0	0	1	4	0	0	0	1.5	0	60
		0	0.5	0	0	0	1.5	0	0	0	0	1.5	0	70
7	*Bacillus licheniformis AJT*	0	4	5	0.5	0	4.5	5	0	0	1	4.5	4.5	40
		0	3	5	0	0	4	5	5	0	0	5.5	5	50
		0	0.5	4.5	0	0	0	5	0	0	0	1.5	0	60
		0	0.5	0	0	0	0.5	0	0	0	0	0	0	70
7	*Bacillus licheniformis AJVa*	0	3	4.5	0	0	4	5	0	0	0	4.5	2.5	40
		0	2	5	0	0	4	5	4.5	0	0	5	3	50
		0	0	1.5	0	0	0	3.5	0	0	0	0.5	0	60
		0	1.5	0	0	0	1	0	0	0	0	0	0	70
7	*Bacillus licheniformis AJVb*	0	4	5	0	0	3.5	5	0	0	0	4	2	40
		0	2	5	0	0	3	5	4.5	0	0	1.5	3	50
		0	3	1.5	0	0	0	3	0	0	0	0	0	60
		0	2.5	0	0	0	1.5	0	0	0	0	0.5	0	70
7	*Brevibacillus agri AJY*	0	0	0	0.5	0	0.5	0	0	0	0	0	0	40
		0	0	0	0	0	0	0	0	0	0	0	0	50
		0	0	0	0	0	0	0	0	0	0	0	0	60
		0	0	0	0	0	0	0	0	0	0	0	0	70
	Unknown sp. AIH 1	0	0	0	0	0	1	1	1	0	2	0	0	40
		0	0	0	0	0	1	1	0	0	2	0	0	50
		0	0	0	0	0	0	0	0	0	0	0	0	60
		0	0	0	0	0	0	0	0	0	0	0	0	70
	Unknown sp. AIS1b	0	0	0	0	0	1	0	0	0	0	0	0	40
		0	1	0	0	0	1	0	0	0	0	0	0	50

	0	2	1	0	0	1.5	0	1	0	0.5	0	0	60
	0	1.5	1.5	0	0	2	0	0	0	1	0	0	70
Unknown sp. AIU	0	0	0	0	0	0	0	0	0	0	0	0	40
	0	0	1	0	0	0	0	0	0	0	0	0	50
	0	0	0	1	0	0	0	0	0	0	0	0	60
	0	0	0	0	0	0	0	0	0	0	0	0	70
Unknown sp. AIWa	0	0	0	0	0	0	0	0	0	0	0	0	40
	0	0.5	0	0	0	1	0	0	0	0	0	0	50
	0	1	0	0	0	1	0	0	0	0	0	0	60
	0	1	0	0	0	1	0	0	0	0	0	0	70
Unknown sp. AJP	0	0	0	0	0	0	0	0	0	0	0	0	40
	0	0	0	0	0	0	0	0	0	0	0	0	50
	0	0	0	0	0	0	0	0	0	0	0	0	60
	0	0	0	0	0	0	0	0	0	0	0	0	70
Unknown sp. AKE	0	0	0	0	0	0	0	0	0	0	0	0	40
	0	1.5	0	0	0	1	0	0	0	0	0	0	50
	0	0	0	0	0	0	0	0	0	0	0	0	60
	0	0	0	0	0	0	0	0	0	0	0	0	70

8.1.5 Plate morphology

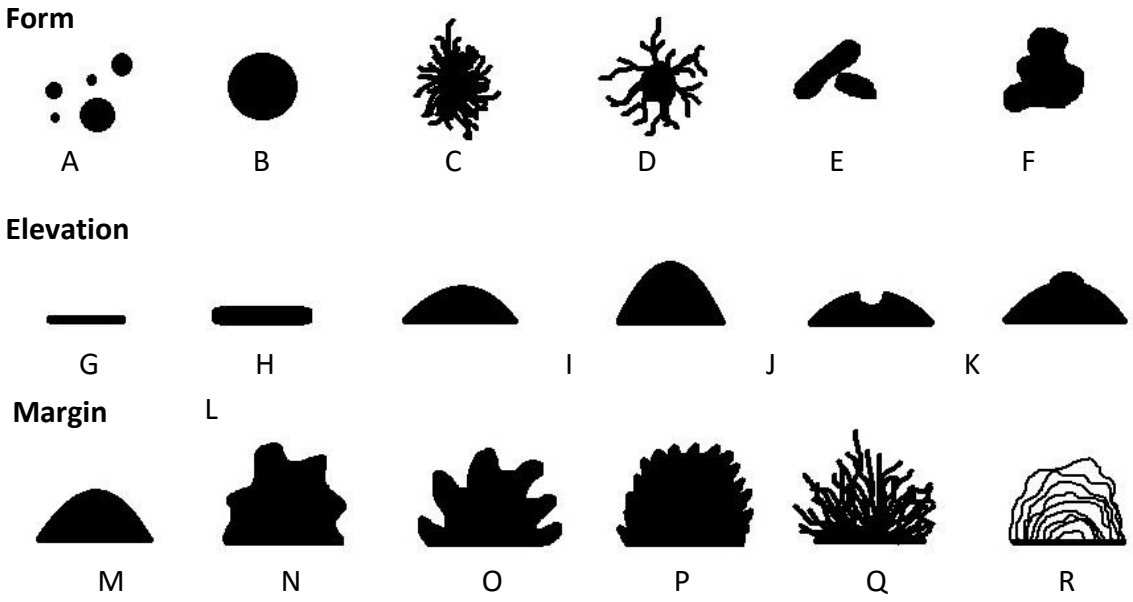


Figure 8.1 Colony morphology of bacteria on agar plates.

A – punctiform, B – circular, C – filamentous, D – rhizoid, E – spindle, F – irregular, G – flat, H – raised, I – convex, J – pulvinate, K – crateriform, L – umbonate, M – entire, N – undulate, O – lobate, P – erose, Q – filamentous, R – curled. Based on (Toth *et al.*, 2013).

8.1.6 Morphology comparison of thermotolerant isolates grown on Nutrient and MRS agar

Thermophilic strains were isolated on both Nutrient and MRS agar. Single colonies could not always be obtained from MRS agar due to the high growth rate overnight and potential production of extracellular polysaccharides (Figure 8.2). Growth appeared restricted on Nutrient agar and single colonies could be obtained more easily (Figure 8.3).

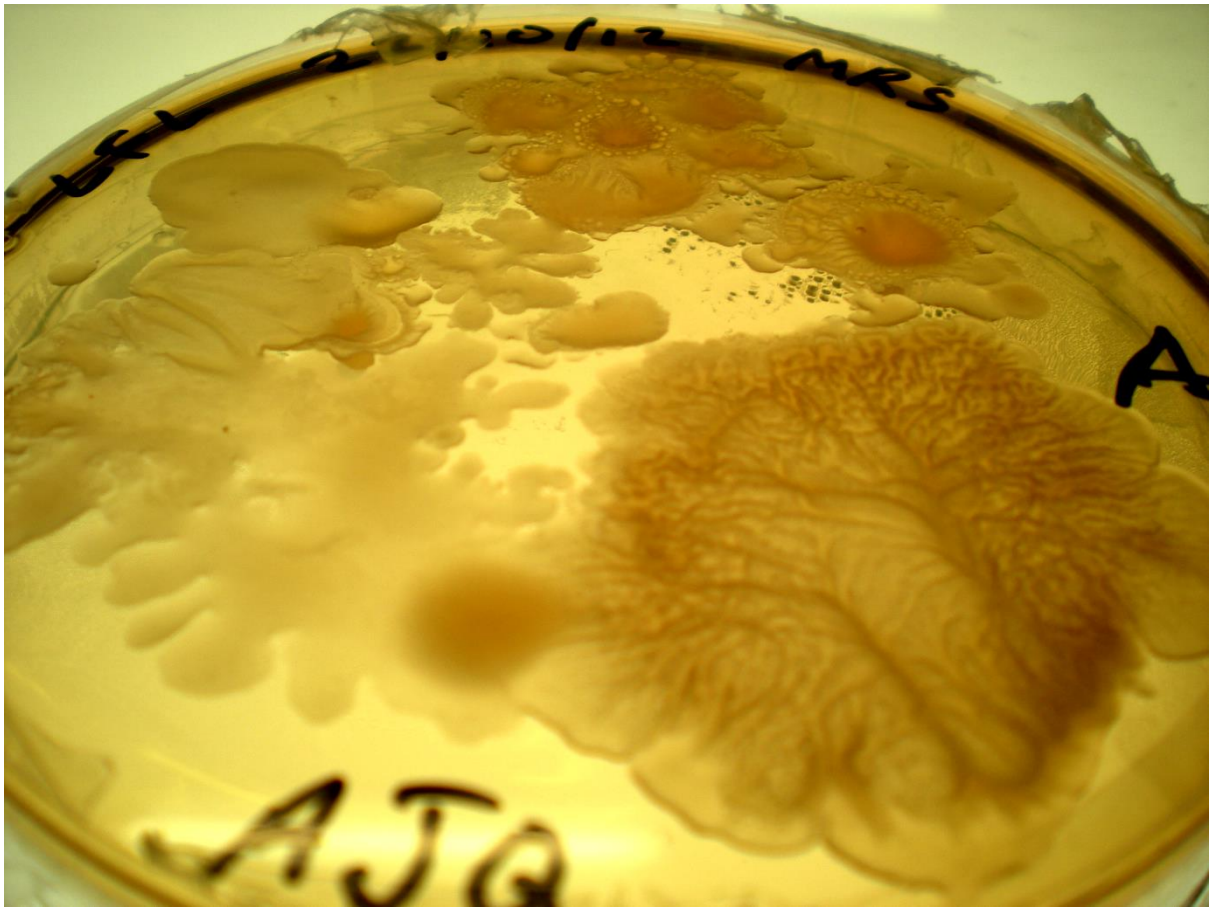


Figure 8.2 Morphology of thermophilic bacterial isolate from *Miscanthus* chip grown on MRS agar overnight.

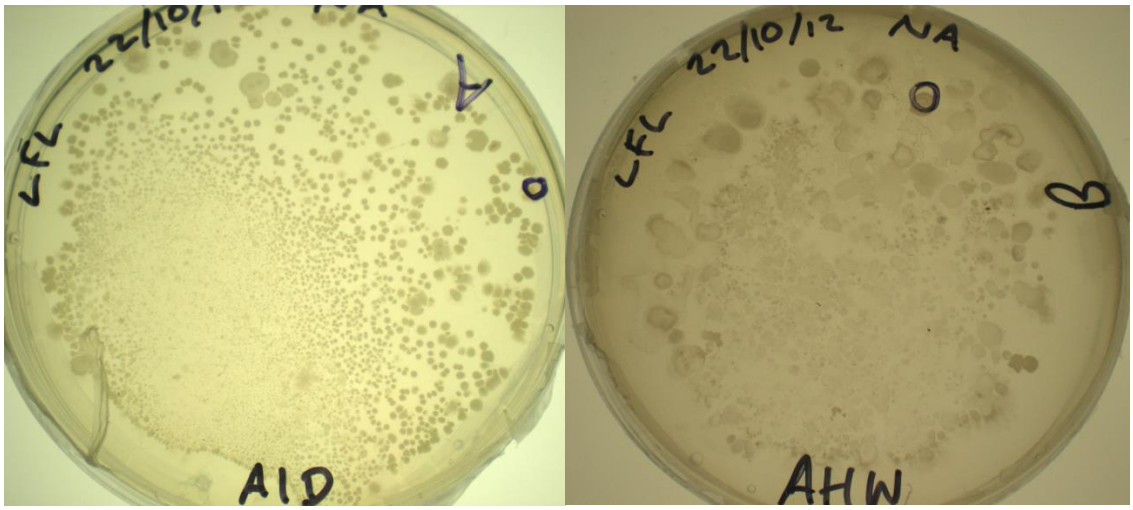


Figure 8.3 Morphology of thermotolerant bacterial isolates from *Miscanthus* chip grown on Nutrient agar overnight.

8.2 Appendix 2: Enzyme Isolation and Expression

8.2.1 Phadebas Amylase Test Standard Curve

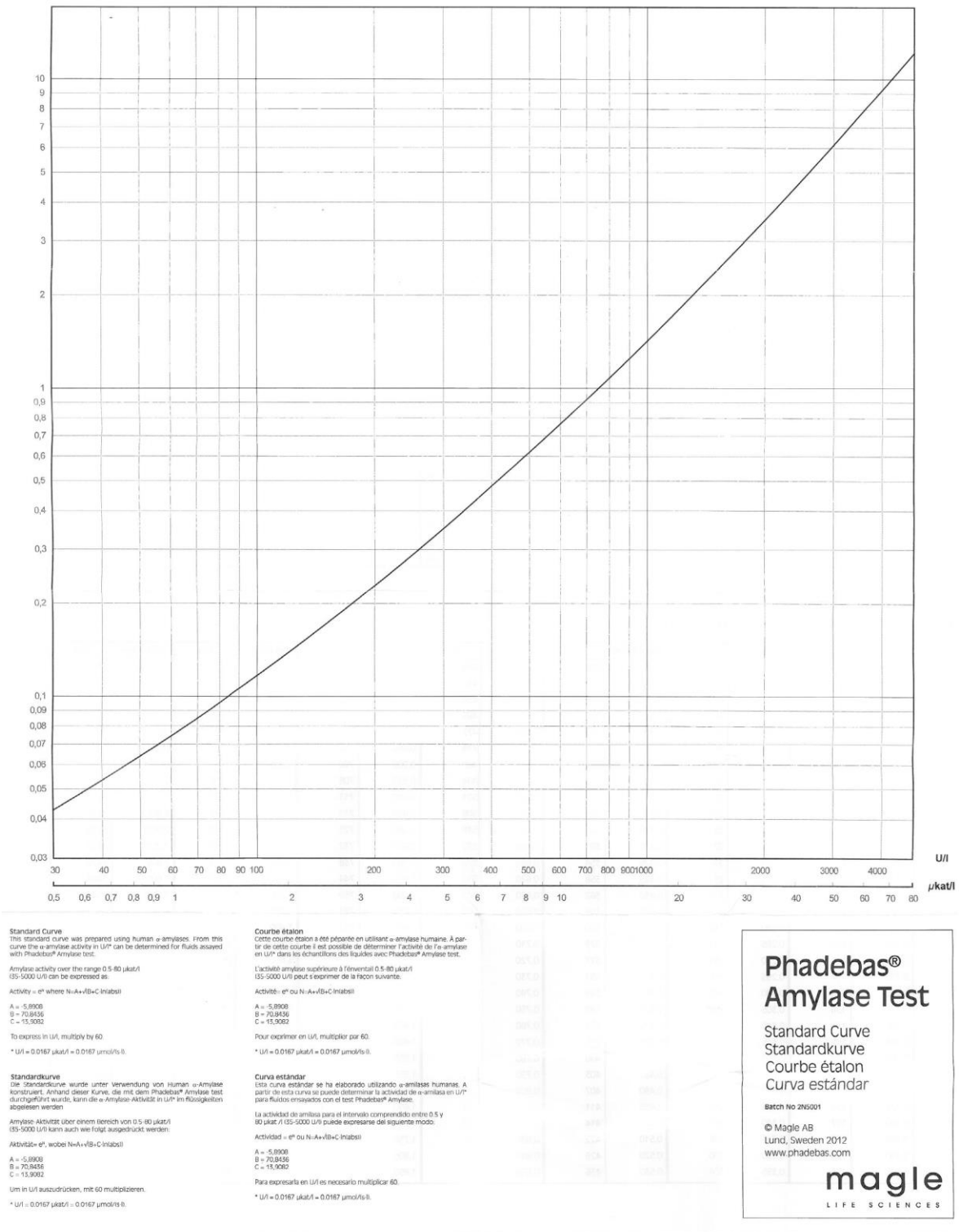


Figure 8.4 Phadebas Amylase Test (Magle Life Sciences) standard curve for U/l and µkat/l.

8.2.2 AmyA alpha-amylase alignment

Alpha-amylase (AmyA) gene alignment from *B. coagulans* ABQ genome sequence (fig | 345219.18); *B. coagulans* ABQ expressed protein nucleotide sequence (BL21star_expressed_alpha); *B. coagulans* 36D1 genome (alpha-amylase_36D1); *B. coagulans* 2-6 genome (alpha-amylase_2-6), and alternative AmyA alpha-amylase from *B. coagulans* ABQ genome. ClustalW2 was used for multiple alignment of amylase genes (EMBL-EBI, 2014).

```
CLUSTAL 2.1 multiple sequence alignment


fig|345219.18.peg.2080_Cytopla      -------------------------------------MERNHTIMQFFEW 13
BL21star_expressed_alpha-amyla      -------------------------------------MERNHTIMQFFEW 13
alpha_amylase_36D1                  -------------------------------------LERNHTILQFFEW 13
alpha_amylase_2-6                   ---------------------------------------------MQFFEW 6
fig|345219.18.peg.857_Cytoplas      MEKKFFSRLSILMLSLLLVAGSISYFPKSAKAYTSGTSLDNRVIFQSFSL 50
                                                                              :* *.

fig|345219.18.peg.2080_Cytopla      NTPADGSHWNRLKEMAPELKKSGIDAVWLPPVTK--GQSDMDNGYGVYDH 61
BL21star_expressed_alpha-amyla      NTPADGSHWNRLKEMAPELKKSGIDAVWLPPVTK--GQSDMDNGYGVYDH 61
alpha_amylase_36D1                  NTPADGSHWNRLKEMAPELKKTGIDAVWLPPVTK--GQSDMDNGYGVYDH 61
alpha_amylase_2-6                   NTPADGSHWNRLKEMAPELKKTGIDAVWLPPVTK--GQSDMDNGYGVYDH 54
fig|345219.18.peg.857_Cytoplas      YMPYESNMYKILSAKGSELKDWGITDIWLPPAYRSFNMARYMEGYAIADR 100
                                      * :.. :: *.  ..***. **  :****. :  . :   :**.: *:

fig|345219.18.peg.2080_Cytopla      YDLGEFDQ--KGTVRTKYGTKQQLHEAINACHEHDIQVYIDVVMNHKAGA 109
BL21star_expressed_alpha-amyla      YDLGEFDQ--KGTVRTKYGTKQQLHEAINACHEHDIQVYIDVVMNHKAGA 109
alpha_amylase_36D1                  YDLGEFDQ--KGTVRTKYGTKQQLHEAINACHEHDIQVYIDVVMNHKAGA 109
alpha_amylase_2-6                   YDLGEFDQ--KGTVRTKYGTKQQLHEAINACHEHDIQVYIDVVMNHKAGA 102
fig|345219.18.peg.857_Cytoplas      YDLGEFNQGPNNTRPTKYGTSDELKSMVSALHASGLKVQEDLVPNQVLGL 150
                                    ******:*  :.*  *****.::*:. :.* *  .::*  *:* *:  *

fig|345219.18.peg.2080_Cytopla      DETESFQVVEVDP------------------------------------ 122
BL21star_expressed_alpha-amyla      DETESFQVVEVDP------------------------------------ 122
alpha_amylase_36D1                  DETESFQVVEVDP------------------------------------ 122
alpha_amylase_2-6                   DETEAFQVVEVDP------------------------------------ 115
fig|345219.18.peg.857_Cytoplas      GKREAVYVTRVDQNGNLFKNPYTTGLTTQIRADLYLAYTKGGGEGQAKYG 200
                                    .: *:. *..**

fig|345219.18.peg.2080_Cytopla      ------------------------------------------------
BL21star_expressed_alpha-amyla      ------------------------------------------------
alpha_amylase_36D1                  ------------------------------------------------
alpha_amylase_2-6                   ------------------------------------------------
fig|345219.18.peg.857_Cytoplas      YIKEWNKKYFNGTSVQGQGMDRVMKDSEGIPYRYFGPNNPKNHLPSWLNE 250


fig|345219.18.peg.2080_Cytopla      ---MDRNKEISEPFEIEGWTKFNFTNRKGKYSD----------FTWNHTH 159
BL21star_expressed_alpha-amyla      ---MDRNKEISEPFEIEGWTKFNFTNRKGKYSD----------FTWNHTH 159
alpha_amylase_36D1                  ---MDRNKEISEPFEIEGWTKFNFTNRKGKYSD----------FTWNHTH 159
alpha_amylase_2-6                   ---MDRTKEISEPFEIEGWTKFNFTNRKDKYSD----------FTWNHTH 152
fig|345219.18.peg.857_Cytoplas      AAAANKINTVDTYFSVDGWYAAKDASTSDNYWKPMLMNYDPGYLKYMKSH 300
                                        :: : :.  *.::**   : :. ..:* .          :.: ::*

fig|345219.18.peg.2080_Cytopla      FSGVDYDNRTGRNG------------------------------------ 173
BL21star_expressed_alpha-amyla      FSGVDYDNRTGRNG------------------------------------ 173
alpha_amylase_36D1                  FSGVDYDNRTGRNG------------------------------------ 173
alpha_amylase_2-6                   FSGVDYDNRTGRNG------------------------------------ 166
fig|345219.18.peg.857_Cytoplas      GYSTVDDILNGDNGEIASLTDAYIASQPGYGFGSEERSFKNDNSGSDDQD 350
                                      ..  *  .* **

fig|345219.18.peg.2080_Cytopla      IFRIVGENKHWSEHVDNEFG-NFDYLMYADIDYNHPDVKKEMIEWGKWLA 222
BL21star_expressed_alpha-amyla      IFRIVGENKHWSEHVDNEFG-NFDYLMYADIDYNHPDVKKEMIEWGKWLA 222
alpha_amylase_36D1                  IFRIVGENKHWSEHVDNEFG-NFDYLMYADIDYNHPDVKKEMIEWGKWLA 222
alpha_amylase_2-6                   IFRIVGENKHWDEHVDNEFG-NFDYLMYADIDYNHPDVKKEMIEWGKWLA 215
```

```
fig|345219.18.peg.857_Cytoplas      QFLFVKKNGTTLHNLNNTISGKKQFLLGMDIDNGNPTVQKEQIHWMNWLL 400
                                     * :* :*    .:::* :. : ::*:  *** .:* *:** *.* :**

fig|345219.18.peg.2080_Cytopla      DTTGCDGYRLDAIKHINHDFIRDFAAALMEHRGDHFYFVGEFWNPQLEAC 272
BL21star_expressed_alpha-amyla      DTTGCDGYRLDAIKHINHDFIRDFAAALMEHRGDHFYFVGEFWNPQLEAC 272
alpha_amylase_36D1                  DTTGCDGYRLDAIKHINHDFIRDFAAALMEHRGDHFYFVGEFWNPQLEAC 272
alpha_amylase_2-6                   DTTGCDGYRLDAIKHINHDFIRDFAAALMEHRGDHFYFVGEFWNPQLEAC 265
fig|345219.18.peg.857_Cytoplas      DTYQFDGFRIDAASHYDQQVLLDEADVMKQHFGNNLNDHLSYIETYESAG 450
                                    **   **:*:** .* :::.: * * .: :* *:::    .: :.  .*

fig|345219.18.peg.2080_Cytopla      QKYLDHVQFKIDLFDVALHYKLHEAS-KKGRAFDLTTIFHDTLVQTHPLN 321
BL21star_expressed_alpha-amyla      QKYLDHVQFKIDLFDVALHYKLHEAS-KKGRAFDLTTIFHDTLVQTHPLN 321
alpha_amylase_36D1                  QKYLDHVQFKIDLFDVALHYKLHEAS-KKGRAFDLPTIFHDTLVQTHPLN 321
alpha_amylase_2-6                   QKYLDHVQFKIDLFDVALHYKLHEAS-KKGRAFDLTTIFHDTLVQTHPLN 314
fig|345219.18.peg.857_Cytoplas      TNFENANGNPQLMMDYALFYSLQNALGKNSPSNNLSTIATNAVVNRAGAG 500
                                     :: :       ::* **.*.*::*  *:. : :*.**  :::*:    .

fig|345219.18.peg.2080_Cytopla      ------AVTFVDNHDSQPN------------------------------- 334
BL21star_expressed_alpha-amyla      ------AVTFVDNHDSQPN------------------------------- 334
alpha_amylase_36D1                  ------AVTFVDNHDSQPN------------------------------- 334
alpha_amylase_2-6                   ------AVTFVDNHDSQPN------------------------------- 327
fig|345219.18.peg.857_Cytoplas      TANATPNWSFVNNHDQEKNRVNSIMLDQYGIKPGTHYGTSTPKAFQDLYD 550
                                           :**:***.: *

fig|345219.18.peg.2080_Cytopla      ------------ESLESWVDDWFKQ---SAYALILLRKDGYPCVFYGDMY 369
BL21star_expressed_alpha-amyla      ------------ESLESWVDDWFKQ---SAYALILLRKDGYPCVFYGDMY 369
alpha_amylase_36D1                  ------------ESLESWVDDWFKQ---SAYALILLRKDGYPCVFYGDMY 369
alpha_amylase_2-6                   ------------ESLESWVDDWFKQ---SAYALILLRKDGYPCVFYGDMY 362
fig|345219.18.peg.857_Cytoplas      KKTEAKALDIYEKDMESTVKKYAPSNVPSQYAYVLTNKDTVPTVFYGDLY 600
                                              :.:** *..:  .   * ** :* .**  * *****:*

fig|345219.18.peg.2080_Cytopla      GIGGDNPIPGK--KDALSPLLSVRREKAYGEQDDYFDHPNTIGWVRRGVP 417
BL21star_expressed_alpha-amyla      GIGGDNPIPGK--KDALSPLLSVRREKAYGEQDDYFDHPNTIGWVRRGVP 417
alpha_amylase_36D1                  GIGGDNPIPGK--KDALSPLLSVRREKAYGEQDDYFDHPNTIGWVRRGVP 417
alpha_amylase_2-6                   GIGGDHPIPGK--KGALSPLLSVRREKAYGEQDDYFDHPNTIGWVRRGVP 410
fig|345219.18.peg.857_Cytoplas      KTNASYMSERTPYYDTIVKLLKVRKNYAYGNQQVTNYKSNTSGTAGKDLI 650
                                      ...     .   .::  **.**:: ***:*:    :.** * . :.:

fig|345219.18.peg.2080_Cytopla      EIP------HSGCAVVISNGENGEK--RMLVGKERAGEVWVDATGNRQEK 459
BL21star_expressed_alpha-amyla      EIP------HSGCAVVISNGENGEK--RMLVGKERAGEVWVDATGNRQEK 459
alpha_amylase_36D1                  EMP------HSGCAVVISNGENGEK--RMLVGKERAGEVWVDATGNRQEK 459
alpha_amylase_2-6                   EMP------HSGCAVVISNGENGEK--RMLAGKERAGEVWVDATGNRQEK 452
fig|345219.18.peg.857_Cytoplas      SSVRYGNDRNTGVATVIGNNPKTDTTIKVNMGSRHANQTFEDATGFHNEK 700
                                    .        ::* *.**.*. : :.  ::  *..:*.:.: **** ::**

fig|345219.18.peg.2080_Cytopla      VTIGEDGYAGFPVNG-------GSVSVWVQETDEN--------------- 487
BL21star_expressed_alpha-amyla      VTIGEDGYAGFPVNG-------GSVSVWVQETGENLEHHHHHH------- 495
alpha_amylase_36D1                  VTIGEDGYAGFPVNG-------GSVSVWVQETDEN--------------- 487
alpha_amylase_2-6                   ITIGEDGYAAFPVNG-------GSVSVWVQETGEN--------------- 480
fig|345219.18.peg.857_Cytoplas      LVTDSKGILTVHVKGTQNARVKGYLGVWIPAKKAATPKQGPALKYGKYVT 750
                                    :. ...*   . *:*       * :.**:  .

fig|345219.18.peg.2080_Cytopla      --------------------------------------------------
BL21star_expressed_alpha-amyla      --------------------------------------------------
alpha_amylase_36D1                  --------------------------------------------------
alpha_amylase_2-6                   --------------------------------------------------
fig|345219.18.peg.857_Cytoplas      VTNKHYAVYQDFNWKKKNVNAVNKTYLAKVQYHHSNGSTYLSLYDGKGKW 800

fig|345219.18.peg.2080_Cytopla      --------------------------------------------------
BL21star_expressed_alpha-amyla      --------------------------------------------------
alpha_amylase_36D1                  --------------------------------------------------
alpha_amylase_2-6                   --------------------------------------------------
fig|345219.18.peg.857_Cytoplas      AGYINAKAAKTGSGKQGAAIQYGKSVKVTSKNYGVYQNFNWKKKNIRAVN 850

fig|345219.18.peg.2080_Cytopla      ----------------------------------------
BL21star_expressed_alpha-amyla      ----------------------------------------
alpha_amylase_36D1                  ----------------------------------------
alpha_amylase_2-6                   ----------------------------------------
fig|345219.18.peg.857_Cytoplas      KTYLAKYIYYHINGLSYLSLYDNKGKWIGYINAKAVKSK 889
```

8.3 Appendix 3: Next Generation Sequencing of Miscanthus Associated Bacteria and Protein Modelling of Sugar Releasing Enzymes

8.3.1 RAST SEED Viewer subsystem overview of *B. coagulans* ABQ

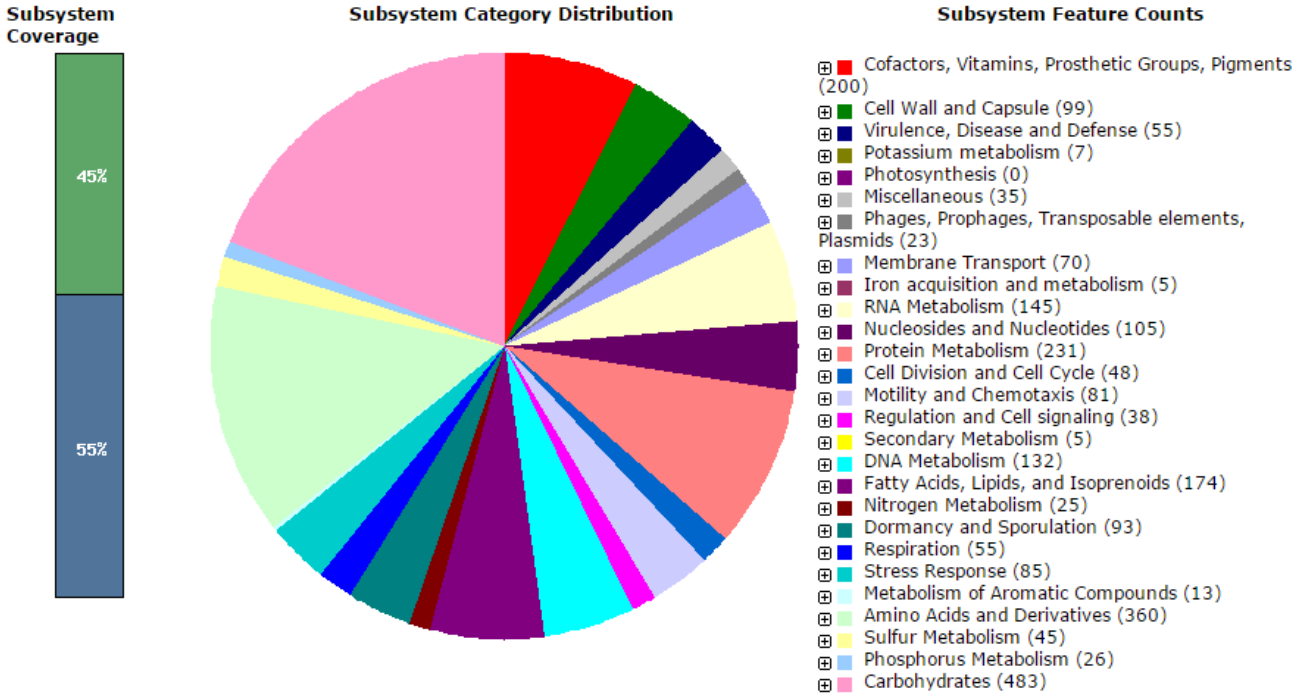


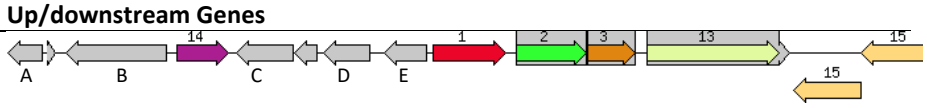
Figure 8.5 Subsystem overview of *Bacillus coagulans* ABQ in the SEED Viewer.

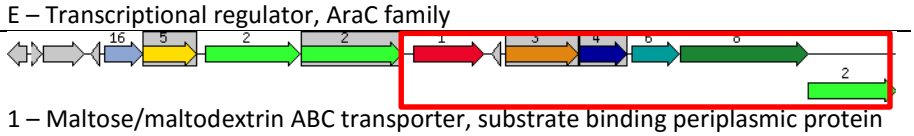
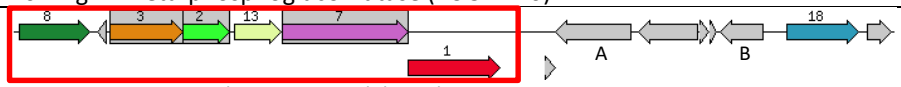
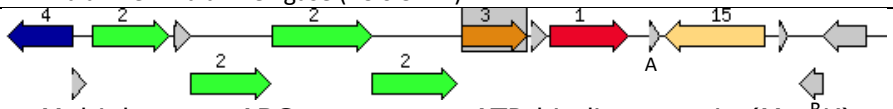
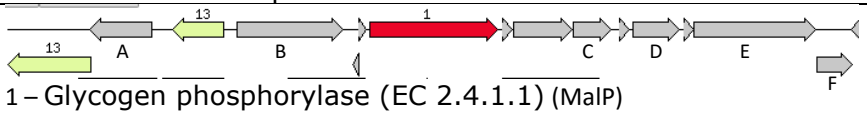
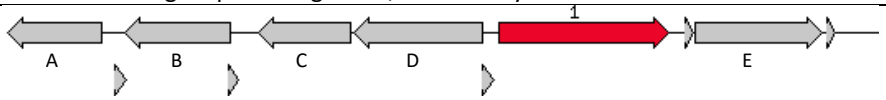
Only 45% of the genome was included in subsystem coverage with 93 hypothetical genes from a total of 1721 identified genes. 2147 genes were not included in subsystem coverage, 1422 of these were classed as hypothetical proteins.

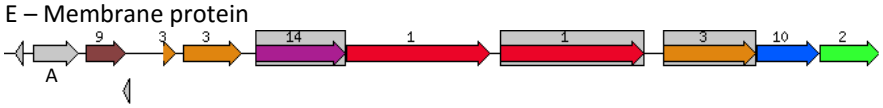
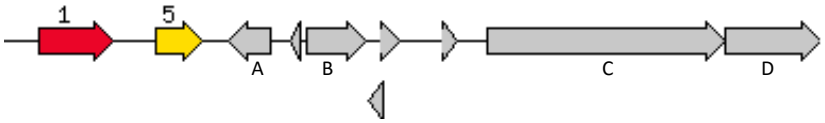
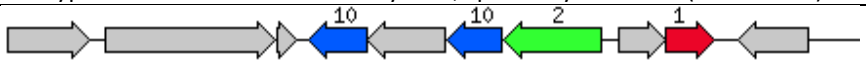
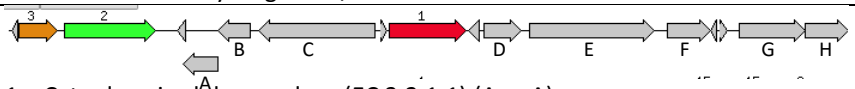
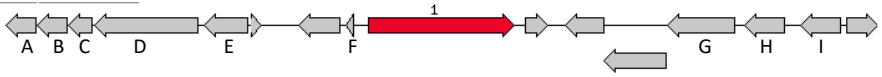
8.3.2 Maltose and maltodextrin utilisation subsystem in *Bacillus coagulans* ABQ

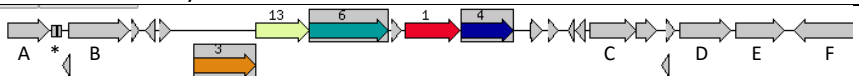
Table 8.6 Maltose and Maltodextrin Utilization subsystem in *Bacillus coagulans* ABQ.

Genes upstream and downstream of each protein involved in the subsystem were recorded. Red box indicated where contig 9 overlapped.

Gene	Contig	Up/downstream Genes
MalE/F/G MapA	33	 1 - Maltose/maltodextrin ABC transporter, permease protein (MalE) 2 – Maltose/maltodextrin ABC transporter, permease protein (MalF) 3 – Maltose/maltodextrin ABC transporter, substrate binding periplasmic protein (MalG) 13 – Maltose phosphorylase (EC 2.4.1.8) / Trehalose phosphorylase (EC 2.4.1.64) (MapA) 14 – Cobalt-zinc-cadmium resistance protein 15 – Cystathionine beta-lyase (EC 4.4.1.8)/Cystathionine gamma-synthase (EC 2.5.1.48) A – Lactocepin (Cell wall-associated serine proteinase) (EC 3.4.21.96)

		B – Phosphomannomutase (EC 5.4.2.8) C – Alcohol dehydrogenase (EC 1.1.1.1) D – N-acetylmannosaminyltransferase (EC 2.4.1.187) E – Transcriptional regulator, AraC family
MalE/F/G/R/A/L AglB PgmB MapA	9	 1 – Maltose/maltodextrin ABC transporter, substrate binding periplasmic protein (MalE) 2 – Neopullulanase (EC 3.2.1.135) (AglB)/Oligo-1,6-glucosidase (EC 3.2.1.10) (MalL) 3 – Maltose/maltodextrin ABC transporter, permease protein (MalF) 4 – Maltose/maltodextrin ABC transporter, permease protein (MalG) 5 – Maltose operon transcriptional repressor (MalR), LacI family 6 – Maltodextrose utilization protein (MalA) 8 – Maltose phosphorylase (EC 2.4.1.8) / Trehalose phosphorylase (EC 2.4.1.64) (MapA) 16 – PgmB Beta-phosphoglucomutase (EC 5.4.2.6)
AglB	9	 1 – Neopullulanase (EC 3.2.1.135) (AlgB) 2 – Maltose/maltodextrin ABC transporter, permease protein (MalG) 3 – Maltose/maltodextrin ABC transporter, permease protein (MalF) 7 – Maltose phosphorylase (EC 2.4.1.8)/Trehalose phosphorylase (EC 2.4.1.64) 8 – Maltose/maltodextrin ABC transporter, substrate binding periplasmic protein (MalE) 13 – Maltodextrose utilization protein (MalA) 18 – Magnesium and cobalt efflux protein CorC A – UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanine ligase (EC 6.3.2.10) B – D-alanine--D-alanine ligase (EC 6.3.2.4)
MsmK	14	 1 – Multiple sugar ABC transporter, ATP-binding protein (MsmK) 2 – Hypothetical protein 3 – Fis-type helix-turn-helix domain protein 4 – Hypothetical protein SAV1846 15 – Fumarate hydratase class II (EC 4.2.1.2) A – Small acid-soluble spore protein, alpha/beta family, SASP_1 B – Mobile element protein
MalP	41	 1 – Glycogen phosphorylase (EC 2.4.1.1) (MalP) 13 – D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) A – Serine-glyoxylate aminotransferase (EC 2.6.1.45) B – Methyl-accepting chemotaxis protein C – Aldehyde dehydrogenase (EC 1.2.1.3) D – Octaprenyl diphosphate synthase (EC 2.5.1.90)/ Dimethylallyltransferase (EC 2.5.1.1)/(2E,6E)-farnesyl diphosphate synthase (EC 2.5.1.10)/Geranylgeranyl diphosphate synthase (EC 2.5.1.29) E – Sensor histidine kinase F – DNA-binding response regulator, LuxR family
MalAZL	70	 1 – Maltodextrin glucosidase (EC 3.2.1.20) (MalZ)

		A – D-beta-hydroxybutyrate permease B – Methylmalonate-semialdehyde dehydrogenase [inositol] (EC 1.2.1.27) C – Alcohol dehydrogenase (EC 1.1.1.1) D – Response regulator of zinc sigma-54-dependent two-component system E – Membrane protein
	13	 1 – Alpha-glucosidase (EC 3.2.1.20) (SusB) 2 – Uronate isomerase (EC 5.3.1.12) 3 – D-mannonate oxidoreductase (EC 1.1.1.57) 9 – 4-Hydroxy-2-oxoglutarate aldolase (EC 4.1.3.16)/ 2-dehydro-3-deoxyphosphogluconate aldolase (EC 4.1.2.14) 10 – Mannonate dehydratase (EC 4.2.1.8) 14 – Oligogalacturonide transporter A – Transcriptional regulator, IclR family
GlvA	1	 1 – Maltose-6'-phosphate glucosidase (EC 3.2.1.122) (GlvA) 5 – PTS system, glucose-specific IIA component (EC 2.7.1.69) A – Glutathione peroxidase family protein B – Transporter, LysE family C – Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72) D – Type I restriction-modification system, specificity subunit S (EC 3.1.21.3)
maa	11	 1 – Maltose O-acetyltransferase (EC 2.3.1.79) (maa) 2 – Osmotically activated L-carnitine/choline ABC transporter, ATP-binding protein OpuCA 10 – Osmotically activated L-carnitine/choline ABC transporter, permease protein OpuCB/OpuCD A – Galactose operon repressor, GalR-LacI family of transcriptional regulators B – Beta-galactosidase (EC 3.2.1.23) C – Osmotically activated L-carnitine/choline ABC transporter, substrate-binding protein OpuCC D – Portion of orf between proU operon and 181 operon E – Short-chain dehydrogenase/reductase SDR
AmyA	39	 1 – Cytoplasmic alpha-amylase (EC 3.2.1.1) (AmyA) 2 – Lipid A export ATP-binding/permease protein MsbA 3 – Deoxyribose-phosphate aldolase (EC 4.1.2.4) A – Putative preQ0 transporter B – NADPH dependent preQ0 reductase (EC 1.7.1.13) C – membrane protein, MmpL family D – Transcriptional regulator, GntR family E – Xylulose-5-phosphate phosphoketolase (EC 4.1.2.9)/Fructose-6-phosphate phosphoketolase (EC 4.1.2.22) F – RsbR, positive regulator of sigma-B G – Carboxynorspermidine dehydrogenase, putative (EC 1.1.1.-) H – Carboxynorspermidine decarboxylase, putative (EC 4.1.1.-)
	24	 1 – Cytoplasmic alpha-amylase (EC 3.2.1.1) (AmyA)

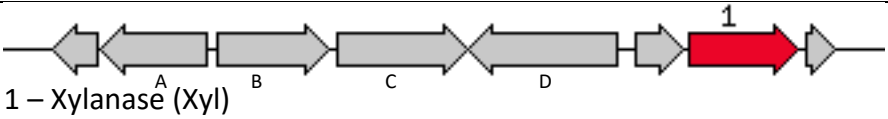
		A – PTS system, glucitol/sorbitol-specific IIB component and second of two IIC components (EC 2.7.1.69) B – PTS system, glucitol/sorbitol-specific IIC component C – Glucitol operon activator protein D – Sorbitol operon transcription regulator E – Sorbitol-6-phosphate 2-dehydrogenase (EC 1.1.1.140) F – putative glucosamine-fructose-6-phosphate aminotransferase G – Modification methylase AvaI (EC 2.1.1.113) (N(4) cytosine-specific methyltransferase AvaI) (M.AvaI) H – Signal transduction histidine kinase I – Putative methyltransferase YodH
AE	8	 1 – Aldose 1-epimerase (EC 5.1.3.3) (AE) 3 – Galactokinase (EC 2.7.1.6) 4 – Galactose operon repressor, GalR-LacI family of transcriptional regulators 6 – Galactose-1-phosphate uridylyltransferase (EC 2.7.7.10) 13 – UDP-glucose 4-epimerase (EC 5.1.3.2) A – Acid phosphatase (EC 3.1.3.2) B – Butyryl-CoA dehydrogenase (EC 1.3.99.2) C – S1 RNA binding domain D – Mannose-6-phosphate isomerase (EC 5.3.1.8) E – Homoserine O-succinyltransferase (EC 2.3.1.46) F – drug resistance transporter, EmrB/QacA family * - tRNA-Ser-CGA/tRNA-Thr-CGT

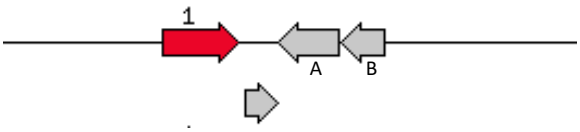
8.3.3 Xylose utilisation pathway in *B. licheniformis* AHY1 and *L. rhizovicinus* AEOb

8.3.3.1 Xylanase genes in L. rhizovicinus AEOb genome

Table 8.7 Xylose utilisation pathway in *Unknown sp.* AEOb.

Xylose utilization using XylAB pathway.

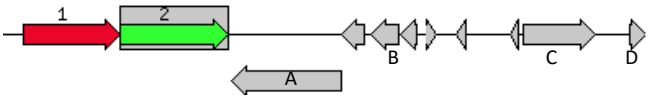
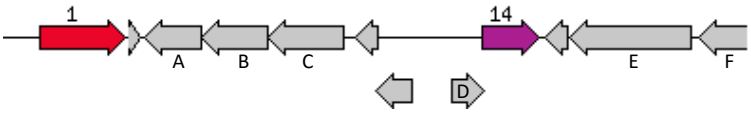
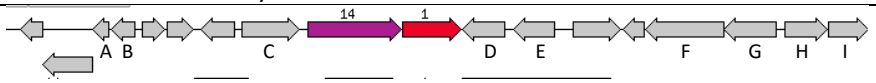
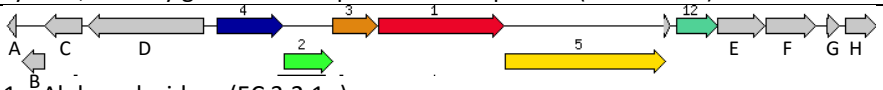
Gene	Contig	Up/downstream Genes
XylA XylR XylB XylG XylF XylH	334	 1 – Xylose isomerase (EC 5.3.1.5) (XylA) 2 – Xylose activator (XylR) (AraC family) 3 – Xylulose kinase (EC 2.7.1.17) (XylB) 4 – D-xylose transport ATP-binding protein (XylG) 5 – Xylose ABC transporter, periplasmic xylose-binding protein (XylF) 6 – Xylose ABC transporter, permease protein (XylH)
Xyl	42	 1 – Xylanase (Xyl) A – Transcriptional regulator, LysR family B – Nucleoside-diphosphate-sugar epimerases C – monooxygenase, FAD-binding D – Manganese transport protein MntH

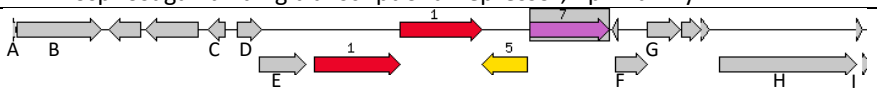
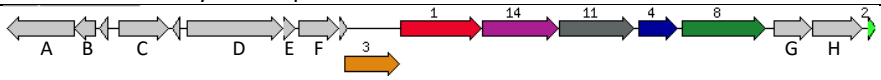
Xyl	489	 1 – Xylanase (Xyl) A – Transcriptional regulator, TetR family B – short-chain dehydrogenase/reductase SDR

8.3.3.2 Xylanase genes in B. licheniformis AHY1 genome

Table 8.8 Xylose utilisation pathway in *Bacillus licheniformis* AHY1.

Xylose utilization using XylAB pathway.

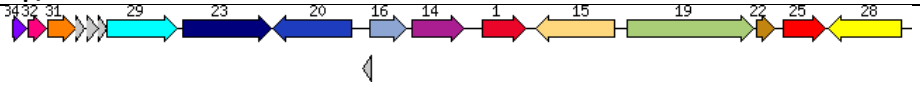
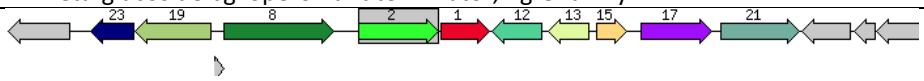
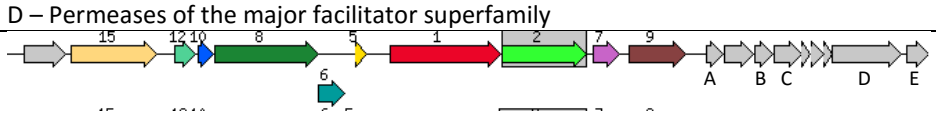
Gene	Contig	Up/downstream Genes
XylA XylB	9	 1 – Xylose isomerase (EC 5.3.1.5) (XylA) 2 – Xylulose kinase (EC 2.7.1.17) (XylB) A – Putative symporter YjcG B – FMN reductase, NADPH-dependent C – UDP-glucose 4-epimerase (EC 5.1.3.2) D – AA3-600 quinol oxidase subunit II
XylR	40	 1 – Xylose-responsive transcription regulator, ROK family (XylR_B S) 14 – CidA-associated membrane protein CidB A – Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3) B – Hydroxyethylthiazole kinase (EC 2.7.1.50) C – LysR family regulatory protein CidR D – Holin-like protein CidA E – Unchracterized transporter YcgO F – LSU m5C1962 methyltransferase RlmI
XylF	15	 1 – Xylose ABC transporter, periplasmic xylose-binding protein (XylF) 14 – methyl-accepting chemotaxis sensory transducer A – Transcriptional regulator, AraC family B – AraC-family transcriptional regulator C – Glycerol dehydrogenase (EC 1.1.1.6) D – CDS_ID OB0544 E – Beta-galactosidase (EC 3.2.1.23) F – S-methylmethionine permease G – Homocysteine S-methyltransferase (EC 2.1.1.10) H – Cellobiose phosphotransferase system YdjC-like protein I – PTS system, N-acetylglucosamine-specific IIB component (EC 2.7.1.69)/ PTS system, N-acetylglucosamine-specific IIC component (EC 2.7.1.69)
bX_ABCa bX_ABCb XylS	23	 1 – Alpha-xylosidase (EC 3.2.1.-) 2 – Possible alpha-xyloside ABC transporter, permease component 3 – Possible alpha-xyloside ABC transporter, permease component

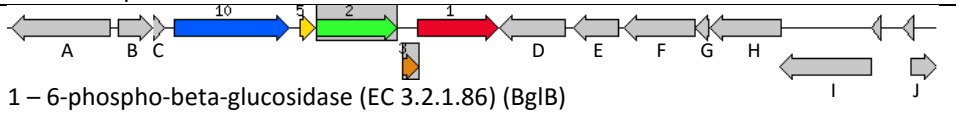
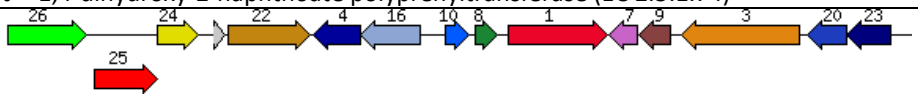
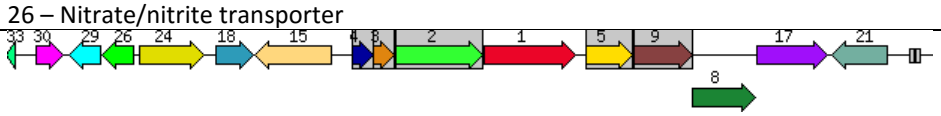
		4 – Possible alpha-xyloside ABC transporter, substrate-binding component 5 – Beta-glucosidase (EC 3.2.1.21) 12 – Manganese ABC transporter, ATP-binding protein SitB A – N-acetylglutamate synthase (EC 2.3.1.1) B – Disulfide bond formation protein BdbC, competence-related C – Periplasmic thiol:disulfide interchange protein DsbA D – Lead, cadmium, zinc and mercury transporting ATPase (EC 3.6.3.3) (EC 3.6.3.5); Copper-translocating P-type ATPase (EC 3.6.3.4) E – Manganese ABC transporter, inner membrane permease protein SitD F – Manganese ABC transporter, periplasmic-binding protein SitA G – Ferredoxin H – Phosphosugar-binding transcriptional repressor, RpiR family
Xyl3	6	 1 – Beta-xylosidase (EC 3.2.1.37) (xyl3) 5 – Transcriptional regulator, AraC family 7 – Lipoprotein A – Transcriptional regulator, RpiR family B – PTS system, maltose and glucose-specific IIC component (EC 2.7.1.69)/ PTS system, maltose and glucose-specific IIB component (EC 2.7.1.69) C – Transcriptional regulator, HxlR family D – Putative membrane protein E – Catechol-2,3-dioxygenase (EC 1.13.11.2) F – NAD(P)H dehydrogenase, quinone family (EC 1.6.99.2) G – Mutator MutT protein H – L-O-lysylphosphatidylglycerol synthase (EC 2.3.2.3) I – Phenazine biosynthesis protein PhzF like
KSAD	15	 1 – Ketoglutarate semialdehyde dehydrogenase (EC 1.2.1.26) (KSAD) 2 – bacitracin ABC transporter, ATP-binding protein 3 – 5-dehydro-4-deoxyglucarate dehydratase (EC 4.2.1.41) 4 – Predicted 5-dehydro-4-deoxyglucarate regulator YcbG 8 – D-galactarate dehydratase (EC 4.2.1.42) 11 – Glucarate dehydratase (EC 4.2.1.40) 14 – D-glucarate permease A – Beta-lactamase regulatory sensor-transducer BlaR1 B – Beta-lactamase repressor BlaI C – Beta-lactamase (EC 3.5.2.6) D – Alkaline phosphatase D precursor (EC 3.1.3.1) E – Twin-arginine translocation protein TatAd F – Twin-arginine translocation protein TatCd G – DNA-binding response regulator H – sensor histidine kinase

8.3.4 Beta-Glucoside Metabolism pathway utilising the Cel(cellobiose) operon for *Bacillus subtilis* AJF2

Table 8.9 Beta-Glucoside Metabolism pathway utilising the Cel(cellobiose) operon for *Bacillus subtilis* AJF2.

Most of the genes are coloured on the contigs as they were homologous to the comparison strain *B. subtilis subsp. subtilis str.* 168 almost perfectly.

Gene	Contig	Up/downstream Genes
BglS BglG	3	 1 – Endo-beta-1,3-1,4 glucanase (Licheninase) (EC 3.2.1.73) (BglS) 14 – Beta-glucoside bgl operon antiterminator, BglG family 15 – Ca(2+) Citrate symporter (TC 2.A.11.1.2) 16 – FIG01230138: hypothetical protein 19 – Catalase (EC 1.11.1.6) 20 – Membrane protein, putative 22 – FIG01231566: hypothetical protein 23 – ATP-dependent RNA helicase YxiN 25 – FIG01242227: hypothetical protein 28 – Nucleoside permease NupC 29 – Rhamnogalacturonan acetylesterase 31 – Uncharacterized protein yxiJ 32 – Uncharacterized protein yxiJ 34 – Hypothetical protein
BglG	236	 1 – Beta-glucoside bgl operon antiterminator, BglG family
BglG LicC	27	 1 – Beta-glucoside bgl operon antiterminator, BglG family 2 – PTS system, sucrose-specific IIB component (EC 2.7.1.69)/ PTS system, sucrose-specific IIC component (EC 2.7.1.69) 8 – FIG01236687: hypothetical protein 12 – Pyridoxamine 5'-phosphate oxidase (EC 1.4.3.5) 13 – YwaF 15 – Putative HTH-type transcriptional regulator ywaE 17 – Tyrosyl-tRNA synthetase (EC 6.1.1.1) 19 – PTS system, cellobiose-specific IIC component (EC 2.7.1.69) (LicC) 21 – Aminopeptidase Y (Arg, Lys, Leu preference) (EC 3.4.11.15) 23 – FIG01227362: hypothetical protein A – Ring canal kelch-like protein B – Putative DNA-binding protein in cluster with Type I restriction-modification system C – Carboxymuconolactone decarboxylase D – Permeases of the major facilitator superfamily
BglF BglB	3	 1 – PTS system, beta-glucoside-specific IIB component (EC 2.7.1.69)/PTS system, beta-glucoside-specific IIC component/PTS system, beta-glucoside-specific IIA component (BglF) 2 – 6-phospho-beta-glucosidase (EC 3.2.1.86) (BglB) 5 – FIG01229148: hypothetical protein 6 – hypothetical protein 7 – Universal stress protein family 8 – FIG01235136: hypothetical protein 9 – Permease of the drug/metabolite transporter (DMT) superfamily 10 – YxiC 12 – FIG01244137: hypothetical protein 15 – FIG01241651: hypothetical protein A – Uncharacterized protein yxzG B – Uncharacterized protein yxiJ C – Uncharacterized protein yxiK

		D – Rhamnogalacturonan acetylesterase E – ATP-dependent RNA helicase YxiN
BglB LicC LicA LicB LicR	27	 1 – 6-phospho-beta-glucosidase (EC 3.2.1.86) (BglB) 2 – PTS system, cellobiose-specific IIC component (EC 2.7.1.69) (LicC) 3 – PTS system, cellobiose-specific IIA component (EC 2.7.1.69) (LicA) 5 – PTS system, lichenan-, cellobiose-specific IIB component (EC 2.7.1.69); PTS system, cellobiose-specific IIB component (EC 2.7.1.69) (LicB) 10 – Transcriptional antiterminator of lichenan operon, BglG family (LicR) A – Catalase (EC 1.11.1.6) B – DNA-3-methyladenine glycosylase II (EC 3.2.2.21) C – Uncharacterized protein yxzF D – Branched-chain amino acid aminotransferase (EC 2.6.1.42) E – DltE F – Poly(glycerophosphate chain) D-alanine transfer protein DltD G – D-alanine--poly(phosphoribitol) ligase subunit 2 (EC 6.1.1.13) H – D-alanyl transfer protein DltB I – D-alanine--poly(phosphoribitol) ligase subunit 1 (EC 6.1.1.13) J – 1,4-dihydroxy-2-naphthoate polyprenyltransferase (EC 2.5.1.74)
BglB LicH	49	 1 – Beta-glucosidase (EC 3.2.1.21) (LicH); 6-phospho-beta-glucosidase (EC 3.2.1.86) (BglB) 3 – methyl-accepting chemotaxis protein 4 – Probable amino-acid ABC transporter permease protein yckA 7 – Nin 8 – YckD 9 – Competence-specific nuclease 10 – FIG01229669: hypothetical protein 16 – Lysine-arginine-ornithine-binding periplasmic protein precursor (TC 3.A.1.3.1) 20 – 6-phospho-3-hexuloisomerase 22 – Putative metal chaperone, involved in Zn homeostasis, GTPase of COG0523 family 23 – D-arabino-3-hexulose 6-phosphate formaldehyde-lyase (EC 4.1.2.43) 24 – ErfK/YbiS/YcfS/YnhG superfamily 25 – GTP cyclohydrolase I (EC 3.5.4.16) type 2 26 – Nitrate/nitrite transporter
LicH LicC LicA LicB	12	 1 – Beta-glucosidase (EC 3.2.1.21) (LicH) 2 – PTS system, cellobiose-specific IIC component (EC 2.7.1.69) (LicC) 3 – PTS system, cellobiose-specific IIA component (EC 2.7.1.69) (LicA) 4 – PTS system, cellobiose-specific IIB component (EC 2.7.1.69) (LicB) 5 – Transcriptional regulator, GntR family 8 – Mannose-6-phosphate isomerase (EC 5.3.1.8) 9 – Fructokinase (EC 2.7.1.4) 15 – L-Proline/Glycine betaine transporter ProP 17 – Mannan endo-1,4-beta-mannosidase precursor (EC 3.2.1.78) 18 – DUF1541 domain-containing protein 21 – catalase, Mn-containing 24 – Deoxyguanosinetriphosphate triphosphohydrolase (EC 3.1.5.1) 26 – acetyltransferase, GNAT family 29 – FIG01228614: hypothetical protein 30 – FIG01239131: hypothetical protein

8.3.5 Search for endo-1,4-beta-xylanases in *B. licheniformis* genomes

Alignment of endo-1,4-beta-xylanases found in two of eight *B. licheniformis* strains with genes of highest homology from NCBI BLASTp search plus top UniProt search for XylY in *Bacillus*.

Table 8.10 Key to ClustalW2 multiple sequence alignment of endo-1,4-beta-xylanases in *B. licheniformis* genomes.

Enzyme gene name	Bacterial isolate genome
fig|6666666.106063.peg.85	*Bacillus licheniformis* AIB
fig|6666666.106073.peg.3353	*Bacillus licheniformis* AJG
fig|6666666.106073.peg.1854	*Bacillus licheniformis* AJG
gi|727789732|ref|WP_033884059	*Bacillus subtilis* (arabinoxylan arabinofuranohydrolase)
gi|510144228|gb|AGN37959.1|	*Bacillus licheniformis* 9945A (arabinoxylan arabinofuranohydrolase XynD)
gi|757434240|gb|AJO17568.1	*Bacillus licheniformis* (Xylanase Y)
tr|O52730|O52730_9BACI	*Bacillus sp.* KK-1 (Xylanase Y (UniProt))

```
gi|727789732|ref|WP_033884059.      MRKKCSVCLWILVLLLSCLSGKSAYAATGTTIAKHIGNSNPLIDHHLGAD 50
fig|6666666.106063.peg.85_Endo      MRKKCSVCLWILVLLLSCLSGKSAYAATSTTIAKHIGNSNPLIDHHLGAD 50
gi|510144228|gb|AGN37959.1|_ar      MKMKCGACLLVLVLFMSCIPGKSVLAAS-TTIAKHVGNSNPLIDHHLGAD 49
fig|6666666.106073.peg.3353_En      MKMKCGACLLVLVLFMSCIPGKSVLAAS-TTIAKHVGNSNPLIDHHLGAD 49
fig|6666666.106073.peg.1854_En      -MRKSLKWIMAVFIGLTCFC-----AAYSQTAMAAVHSKTPDILGTTGKN 44
gi|757434240|gb|AJO17568.1|_xy      -MRKSLKWIMAVFIGLTCFC-----AAYSQTAMAAVHSKTPDILGTTGKN 44
tr|O52730|O52730_9BACI              -MRKSLKWIMAVFIGLTCFC-----TAYSQTAMAAVHSKTPDILGTTGKN 44
                                       *.   :  :.: ::*:      :*   *    : ...* *    * :

gi|727789732|ref|WP_033884059.      PVALTYNGRVYIYMSSDDYEYNSNGTIKDNSFANLNRVFVISSADMVNWT 100
fig|6666666.106063.peg.85_Endo      PVALTYNGRVYIYMSSDDYEYNSNGTIKDNSFANLNRVFVISSADMVNWT 100
gi|510144228|gb|AGN37959.1|_ar      PFALTYNGRVYIYMSSDDYEYNSNGTIKDNSFANLNKVFVISSADMVNWT 99
fig|6666666.106073.peg.3353_En      PFALTYNGRVYIYMSSDDYEYNSNGTIKDNSFANLNKVFVISSADMVNWT 99
fig|6666666.106073.peg.1854_En      NLNQAYK---------KYFDTKGDGK----------------GGSLFHYM 69
gi|757434240|gb|AJO17568.1|_xy      NLNQAYK---------KYFDTKGDGK----------------GGSLFHYM 69
tr|O52730|O52730_9BACI              NLNQAYK---------KYFDTKGDGK----------------GGSLFHYM 69
                                     .  :*:         . :: :.:*.                ...:.::

gi|727789732|ref|WP_033884059.      DHGAIPVAGANGANGGRGIAKWAGASWAPSIAVKKINGKDKFFLYFANSG 150
fig|6666666.106063.peg.85_Endo      DHGAIPVAGANGANGGRGIAKWAGASWAPSIAVKKINGKDKFFLYFANSG 150
gi|510144228|gb|AGN37959.1|_ar      DHGAIPVAGANGANGGRGIAKWAGASWAPSVAVKKINGKDKFFLYFANGG 149
fig|6666666.106073.peg.3353_En      DHGAIPVAGANGANGGRGIAKWAGASWAPSVAVKKINGKDKFFLYFANGG 149
fig|6666666.106073.peg.1854_En      KDGSAYIASTTDDNLGNGYYSVKTEGMSYGMMITLQMNDEYKFQKLWDFV 119
gi|757434240|gb|AJO17568.1|_xy      KDGSAYIASTTDDNLGNGYYSVKTEGMSYGMMITLQMNDEYKFQKLWDFV 119
tr|O52730|O52730_9BACI              KDGSAYIASTTDDNLGNGYYSVKTEGMSYGMMITLQMNDEYKFQKLWDFV 119
                                    ..*:  :*.:.. * *.*  .    . : .: :.   ..:  *  : :

gi|727789732|ref|WP_033884059.      GGIGVLTADSPIGPWTDPIGKPLVTPSTPGMSGVVWLFDPAVFVDDDGTG 200
fig|6666666.106063.peg.85_Endo      GGIGVLTADSPIGPWTDPIGKPLVTPSTPGMSGVVWLFDPAVFVDDDGTG 200
gi|510144228|gb|AGN37959.1|_ar      GGIGVLTADSPTGPWTDPIGKALVTISTPGMSGVVWLFDPAVFVDDDGTG 199
fig|6666666.106073.peg.3353_En      GGIGVLTADSPTGPWTDPIGKALVTISTPGMSGVVWLFDPAVFVDDDGTG 199
fig|6666666.106073.peg.1854_En      RKYMRHSDRSDS-----LYGYHSWHMKTNGSD--VQTIDQNVASDGE--- 159
gi|757434240|gb|AJO17568.1|_xy      RKYMRHSDRSDS-----LYGYHSWHMKTNGSD--VQTIDQNVASDGE--- 159
tr|O52730|O52730_9BACI              RKYMRHSDRNDS-----LYGYHSWHMKTNGSD--VQTIDQNVASDGE--- 159
                                          :  .         *      .* * .  *  :*  *  *.:

gi|727789732|ref|WP_033884059.      YLYAGGGVPGGSNPTQGQWANPKTARVIKLGPDMTSVEGSASTIDAPFMF 250
fig|6666666.106063.peg.85_Endo      YLYAGGGVPGGSNPTQGQWANPKTARVIKLGPDMTSVEGSASTIDAPFMF 250
gi|510144228|gb|AGN37959.1|_ar      YLYAGGGIPGGSNPTQEQWANPKTARVIKLGNDMTSVVGSASVIDAPFMF 249
fig|6666666.106073.peg.3353_En      YLYAGGGIPGGSNPTQEQWANPKTARVIKLGNDMTSVVGSASVIDAPFMF 249
fig|6666666.106073.peg.1854_En      VWFAAALMMASGRWGDKKYPYDYKERAQDMLDALAGDGEYANTGKESRVF 209
gi|757434240|gb|AJO17568.1|_xy      VWFAAALMMASGRWGDKKYPYDYKERAQDMLDALAGDGEYANTGKESRVF 209
tr|O52730|O52730_9BACI              VWFAAALMMASGRWGDKKYPYDYKARAQDMLDALAGDGEYANTGKESRVF 209
                                      :*.. : ....  : ::.   . *. .:   ::.    *.. . . :*

gi|727789732|ref|WP_033884059.      EDSGLHKYNGTYYYSYCINFGGMHPADKPPGEIGYMTSSNPMGPFTYRGH 300
fig|6666666.106063.peg.85_Endo      EDSGLHKYNGTYYYSYCINFGGMHPADKPPGEIGYMTSSNPMGPFTYRGH 300
gi|510144228|gb|AGN37959.1|_ar      EDSGIHKHNGKYYYSYCINFSGVHPANIPPGEIGYMTSSSPMGPFTYRGH 299
fig|6666666.106073.peg.3353_En      EDSGIHKHNGKYYYSYCINFSGVHPANIPPGEIGYMTSSSPMGPFTYRGH 299
fig|6666666.106073.peg.1854_En      IKNSKDQRYAMVRFGPYVNWT--DPSYHVPAFFELFAKSARSSQQYFWKD 257
gi|757434240|gb|AJO17568.1|_xy      IKNSKDQRYAMVRFGPYVNWT--DPSYHVPAFFELFAKSARSSQQYFWKD 257
tr|O52730|O52730_9BACI              IKNSKDQRYAMVRFGPYVNWT--DPSYHVPAFFELFAKSAKSSQQYFWKD 257
                                     ... .:  .   :.  :*:   .*:   *. :  ::.*   .   :  .
```

```
gi|727789732|ref|WP_033884059.      FLKNPGAFFGGGGNNHHAVFNFKNEWYVVYHAQTVSSALFGAGKGYRSPH 350
fig|6666666.106063.peg.85_Endo      FLKNPGAFFGGGGNNHHAVFNFKNEWYVVYHAQTVSSALFGAGKGYRSPH 350
gi|510144228|gb|AGN37959.1|_ar      FLKNPGAFFGAGGNNHHAVFNFKNEWYVVYHAQTVSFALFGAGKGYRSPH 349
fig|6666666.106073.peg.3353_En      FLKNPGAFFGAGGNNHHAVFNFKNEWYVVYHAQTVSFALFGAGKGYRSPH 349
fig|6666666.106073.peg.1854_En      AANKSRTYLSET--------TFKSVLNNGSTVTNAATGLFPDEAGFDG-- 297
gi|757434240|gb|AJO17568.1|_xy      AANKSRTYLSET--------TFKSVLNNGSTVTNAATGLFPDEAGFDG-- 297
tr|O52730|O52730_9BACI              AANKSRTYLSET--------TFKSVLNNGSTVTNAATGLFPDEAGFDG-- 297
                                      ::. :::.          .**.       . ..: .**    *: .

gi|727789732|ref|WP_033884059.      INKLVHNPDGSIQEVAANYAGVTQLSNLNPYNRVEVETFAWNGRILTEKS 400
fig|6666666.106063.peg.85_Endo      INKLVHNPDGSIQEVAANYAGVTQLSNLNPYNRVEAETFAWNGRILTEKS 400
gi|510144228|gb|AGN37959.1|_ar      INKLVHLANGSIQEVAANYRGVSQLANLNPYNRVEAETFAWNGRILTEKS 399
fig|6666666.106073.peg.3353_En      INKLVHLANGSIQEVAANYRGVSQLANLNPYNRVEAETFAWNGRILTEKS 399
fig|6666666.106073.peg.1854_En      ----VSDAAHSSTETDRNFS----YDAWRTVSHIAMDHTLWSS--ADNAY 337
gi|757434240|gb|AJO17568.1|_xy      ----VSDAAHSSTETDRNFS----YDAWRTVSHIAMDHTLWSS--ADNAY 337
tr|O52730|O52730_9BACI              ----VSDAAHSSTETDRNFS----YDAWRTVSHIAMDHTLWSS--ADNAY 337
                                        *  .  *  *.  *:         .. .::  :   *..    :

gi|727789732|ref|WP_033884059.      TAPGGPVNNQHVTSIQNGDWIAVGNADFGAGGARTFKANVASTLGGKIEV 450
fig|6666666.106063.peg.85_Endo      TAPGGPVNNQHVTSIQNGDWIAVGNADFGAGGARTFKANVASTLGGKIEV 450
gi|510144228|gb|AGN37959.1|_ar      SAPGGPENNQNVTSIHNGDWIAVGNADFGTRGARTFKANVASTAGGKIEV 449
fig|6666666.106073.peg.3353_En      SAPGGPENNQNVTSIHNGDWIAVGNADFGTRGARTFKANVASTAGGKIEV 449
fig|6666666.106073.peg.1854_En      RASEQKAVNKFLTFMKR--------ENYGRTAHEYTLNGTAVKKGSPVGL 379
gi|757434240|gb|AJO17568.1|_xy      RASEQKAVNKFLTFMKR--------ENYGRTAHEYTLNGTAVKKGSPVGL 379
tr|O52730|O52730_9BACI              RASEQKAVNKFLTFMKR--------ENYGRTAHEYTLNGTAVKKGSPVGL 379
                                     *.     *: :* ::.          ::*  . .    ..* . *. : :

gi|727789732|ref|WP_033884059.      RLDSADGKLVGTLNVPSTGGAQTWREIETAVSGATGVHKVFFVFTGTGTG 500
fig|6666666.106063.peg.85_Endo      RLDSADGKLVGTLNVPSTGGAQTWREIETEVSGATGVHKVFFVFTGTGTG 500
gi|510144228|gb|AGN37959.1|_ar      RLDSVTGKLVGTLNVPSTGGAQTWREIETTVSGATGVHNVYFVFTGNSSG 499
fig|6666666.106073.peg.3353_En      RLDSVTGKLVGTLNVPSTGGAQTWREIETTVSGATGVHNVYFVFTGNSSG 499
fig|6666666.106073.peg.1854_En      IAANAGGATAASDASLRTGFANAFNSTYIPEDYYGSCLYMLNSLVANGKF 429
gi|757434240|gb|AJO17568.1|_xy      IAANAGGATAASDASLRTGFANAFNSTYIPEDYYGSCLYMLNSLVANGKF 429
tr|O52730|O52730_9BACI              IAANAGGATAASDASLRTGFANAFNSTYIPEDYYGSCLYMLNSLVANGKF 429
                                       .. *  ..:     ** *::::..     .   .   :   :.....

gi|727789732|ref|WP_033884059.      NLLNFDYWQFTQR--- 513
fig|6666666.106063.peg.85_Endo      NLFNFDYWQFTQR--- 513
gi|510144228|gb|AGN37959.1|_ar      NLFNFDYWQFTQNEHR 515
fig|6666666.106073.peg.3353_En      NLFNFDYWQFTQNEHR 515
fig|6666666.106073.peg.1854_En      TMYLP----------- 434
gi|757434240|gb|AJO17568.1|_xy      TMYLP----------- 434
tr|O52730|O52730_9BACI              AMYLP----------- 434
                                     :
```

8.3.6 Beta-glucanase alignment for *B. subtilis* AJF2 endo-beta-1,3-1,4-glucanase (EC 3.2.1.73) verification

Table 8.11 Key to ClustalW2 multiple sequence alignment for *B. subtilis* AJF2 endo-beta-1,3-1,4-glucanase (EC 3.2.1.73).

Identifier	Enzyme gene and bacterial source
gi|489337319	beta-glucanase *Bacillus subtilis* (WP_003244531.1)
fig|936156.11	fig|936156.11.peg.2146 Endo-beta-1,3-1,4-glucanase (Licheninase) (EC_3.2.1.73) *Bacillus subtilis* BSn5
gi|639241500	beta-glucanase *Bacillus subtilis* (WP_024571825.1)
gi|742327606	beta-glucanase *Bacillus subtilis* (WP_038828911.1)
gi|633605343	cellulose *Bacillus subtilis* (gb AHZ57098.1)
fig|6666666	fig|6666666.106072.peg.2457 Endo-beta-1,3-1,4 glucanase (Licheninase) (EC_3.2.1.73) *Bacillus subtilis* AJF2
gi|505296801	beta-glucanase *Bacillus subtilis* (WP_015483903.1)
gi|749034484	beta-glucanase *Bacillus sp.* A053 (WP_040081344.1)
gi|504291323	beta-glucanase *Bacillus subtilis* (WP_014478425.1)

```
gi|489337319|ref|WP_003244531.      MPYLKRVLLLLVTGLFMSLFAVTATASAQTGGSFFDPFNGYNSGFWQKAD 50
fig|936156.11.peg.2146_Endo-be      MPYLKRVLLLLVTGLFMSLFAVTATASAQTGGSFFDPFNGYNSGFWQKAD 50
gi|639241500|ref|WP_024571825.      MPYLKRVLLLLVTGLFMSLFAVTSTASAQTGGSFFDPFNGYNSGFWQKAD 50
gi|742327606|ref|WP_038828911.      MPYLKRVLLLLVTGLFMSLFAVTSTASAQTGESFFDPFNGYNSGFWQKAD 50
gi|633605343|gb|AHZ57098.1| ce      MPYLKRVLLLLVTGLFMSLFAVTSTASAQTGGSFFDPFNGYNSGFWQKAD 50
fig|6666666.106072.peg.2457_En      MPYLKRVLLLLVTGLFMSLFAVTSTASAQTGGSFFDPFNGYNSGFWQKAD 50
gi|505296801|ref|WP_015483903.      MPYLKRVLLLLVTGLFMSLFAVTSTASAQTGGSFFDPFNGYNSGFWQKAD 50
gi|749034484|ref|WP_040081344.      MPYLKRVLLLLVTGLFMSLFAVTSTASAQTGGSFFDPFNGYNSGFWQKAD 50
gi|504291323|ref|WP_014478425.      MPYLKRVLLLLVTGLFMSLFAVTSTASAQTGGSFFDPFNGYNSGFWQKAD 50
```

```
                                    ***********************:******* *****************

gi|489337319|ref|WP_003244531.      GYSNGNMFNCTWRANNVSMTSLGEMRLALTSPAYNKFDCGENRSVQTYGY 100
fig|936156.11.peg.2146 Endo-be      GYSNGNMFNCTWRANNVSMTSLGEMRLALTSPAYNKFDCGENRSVQTYGY 100
gi|639241500|ref|WP_024571825.      GYSNGNMFNCTWRANNVSMTSLGEMRLALTSPAYNKFDCGENRSVQTYGY 100
gi|742327606|ref|WP_038828911.      GYSNGNMFNCTWRANNVSMTSLGEMRLALTSPAYNKFDCGENRSVQTYGY 100
gi|633605343|gb|AHZ57098.1|_ce      GYSNGNMFNCTWRANNVSMTSLGEMRLALTSPAYNKFDCGENRSVQTYGY 100
fig|6666666.106072.peg.2457_En      GYSNGNMFNCTWRANNVSMTSLGEMRLALTSPAYNKFDCGENRSVQTYGY 100
gi|505296801|ref|WP_015483903.      GYSNGNMFNCTWRANNVSMTSLGEMRLALTSPSYNKFDCGENRSVQTYGY 100
gi|749034484|ref|WP 040081344.      GYSNGDMFNCTWRANNVSMTSLGEMRLALTSPSYNKFDCGENRSVQTYGY 100
gi|504291323|ref|WP 014478425.      GYSNGNMFNCTWRANNVSMTSLGEMRLALTSPSYNKFDCGENRSVQTYSY 100
                                    *****:**************************:***************.*

gi|489337319|ref|WP_003244531.      GLYEVRMKPAKNTGIVSSFFTYTGPTDGTPWDEIDIEFLGKDTTKVQFNY 150
fig|936156.11.peg.2146_Endo-be      GLYEVRMKPAKNTGIVSSFFTYTGPTDGTPWDEIDIEFLGKDTTKVQFNY 150
gi|639241500|ref|WP_024571825.      GLYEVRMKPAKNTGIVSSFFTYTGPTDGTPWDEIDIEFLGKDTTKVQFNY 150
gi|742327606|ref|WP_038828911.      GLYEVRMKPAKNTGIVSSFFTYTGPTDGTPWDEIDIEFLGKDTTKVQFNY 150
gi|633605343|gb|AHZ57098.1| ce      GLYEVRMKPAKNTGIVSSFFTYTGPTDGTPWDEIDIEFLGKDTTKVQFNY 150
fig|6666666.106072.peg.2457_En      GLYEVRMKPAKNTGIVSSFFTYTGPTDGTPWDEIDIEFLGKDTTKVQFNY 150
gi|505296801|ref|WP_015483903.      GLYEVRMKPAKNTGIVSSFFTYTGPTDGTPWDEIDIEFLGKDTTKVQFNY 150
gi|749034484|ref|WP_040081344.      GLYEVRMKPAKNTGIVSSFFTYTGPTDGTPWDEIDIEFLGKDTTKVQFNY 150
gi|504291323|ref|WP 014478425.      GLYEVRMKPAKNTGIVSSFFTYTGPTDGTPWDEIDIEFLGKDTTKVQFNY 150
                                    **************************************************

gi|489337319|ref|WP_003244531.      YTNGAGNHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQLKHTATNQI 200
fig|936156.11.peg.2146_Endo-be      YTNGAGNHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQLKHTATNQI 200
gi|639241500|ref|WP_024571825.      YTNGAGNHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQLKHTATNQI 200
gi|742327606|ref|WP_038828911.      YTNGAGNHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQLKHTATNQI 200
gi|633605343|gb|AHZ57098.1|_ce      YTNGAGNHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQLKHTATDQI 200
fig|6666666.106072.peg.2457 En      YTNGAGNHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQLKHTATNQI 200
gi|505296801|ref|WP_015483903.      YTNGAGNHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQLKHTATNQI 200
gi|749034484|ref|WP_040081344.      YTNGAGNHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQLKHTATNQI 200
gi|504291323|ref|WP_014478425.      YTNGAGNHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQLKHTATNQI 200
                                    **************************************************:**

gi|489337319|ref|WP 003244531.      PTTPGKIMMNLWNGTGVDEWLGSYNGVNPLYAHYDWVRYTKK 242
fig|936156.11.peg.2146 Endo-be      PTTPGKIMMNLWNGTGVDEWLGSYNGVNPLYAHYDWVRYTKK 242
gi|639241500|ref|WP_024571825.      PTTPGKIMMNLWNGTGVDEWLGSYNGVNPLYAHYDWVRYTKK 242
gi|742327606|ref|WP_038828911.      PTTPGKIMMNLWNGTGVDEWLGSYNGVNPLYAHYDWVRYTKK 242
gi|633605343|gb|AHZ57098.1|_ce      PTTPGKIMMNLWNGTGVDEWLGSYNGVNPLYAHYDWVRYTKK 242
fig|6666666.106072.peg.2457_En      PTTPGKIMMNLWNGTGVDEWLGSYNGVNPLYAHYDWVRYTKK 242
gi|505296801|ref|WP_015483903.      PTTPGKIMMNLWNGTGVDEWLGSYNGVNPLYAHYDWVRYTKK 242
gi|749034484|ref|WP 040081344.      PTTPGKIMMNLWNGTGVDEWLGSYNGVNPLYAHYDWVRYTKK 242
gi|504291323|ref|WP_014478425.      PTTPGKIMMNLWNGTGVDEWLGSYNGVNPLYAHYDWVRYTKK 242
                                    ******************************************
```

8.3.7 Beta-glucanase alignment to identify KEGG-KASS endo-beta-glucanase in *Bacillus subtilis* AJF2 (fig|6666666.106072.peg.3365)

```
fig|6666666.106072.peg.3365_Ba      MKRSISIFITCLLITLLTMGGIMASPASAAGTKTPVAKNGQLSIKGTQLV 50
M2UBG4|cellulase|EC                 MKRSISIFITCLLITLLTMGGMIASPASAAGTKTPVAKNGQLSIKGTQLV 50
H6WUK0|cellulase|EC                 MKRSISIFITCLLIAVLTMGGLLPSPASAAGTKTPVAKNGQLSIKGTQHV 50
                                    **************::*****:::.************************ *

fig|6666666.106072.peg.3365_Ba      NRDGKAVQLKGISSHGLQWYGEYVNKDSLKWLRDDWGITVFRAAMYTADG 100
M2UBG4|cellulase|EC                 NRDGKAVQLKGISSHGLQWYGEYVNKDSLKWLRDDWGITVFRAAMYTADG 100
H6WUK0|cellulase|EC                 NRDGKTVQLKGISSHGLQWYGDFVNKDSLKWLRDDWGITVFRAAMYTADG 100
                                    *****:***************::*************************

fig|6666666.106072.peg.3365_Ba      GYIDNPSVKNKVKEAVEAAKELGIYVIIDWHILNDGNPNQNKEKAKEFFK 150
M2UBG4|cellulase|EC                 GYIDNPSVKNKVKEAVEAAKELGIYVIIDWHILNDGNPNQNKEKAKEFFK 150
H6WUK0|cellulase|EC                 GYIDNPSVKNKVKEAVEAAKELGIYVIIDWHILNDGNPNQNKEKAKEFFK 150
                                    **************************************************

fig|6666666.106072.peg.3365_Ba      EMSSLYGNTPNVIYEIANEPNGDVNWKRDIKPYAEEVISVIRKNDPDNII 200
M2UBG4|cellulase|EC                 EMSSLYGNTPNVIYEIANEPNGDVNWKRDIKPYAEEVISVIRKNDPDNII 200
H6WUK0|cellulase|EC                 EMSSLYGNTPNVIYEIANEPNGDVNWKRDIKPYAEEVISVIRKNDPDNII 200
                                    **************************************************

fig|6666666.106072.peg.3365_Ba      IVGTGTWSQDVNDAADDQLKDANVMYALHFYAGTHGQFLRDKANYALSKG 250
M2UBG4|cellulase|EC                 IVGTGTWSQDVNDAADDQLKDANVMYALHFYAGTHGQFLRDKANYALSKG 250
H6WUK0|cellulase|EC                 IVGTGTWSQDVNDAADDQLKDANVMYALHFYAGTHGQSLRDKANYALSKG 250
                                    ************************************ ************

fig|6666666.106072.peg.3365_Ba      APIFVTEWGTSDASGNGGVFLDQSREWLKYLDSKTISWVNWNLSDKQESS 300
M2UBG4|cellulase|EC                 APIFVTEWGTSDASGNGGVFLDQSREWLKYLDSKTISWVNWNLSDKQESS 300
H6WUK0|cellulase|EC                 APIFVTEWGTSDASGNGGVFLDQSREWLNYLDSKNISWVNWNLSDKQESS 300
                                    ****************************:*****.***************

fig|6666666.106072.peg.3365_Ba      SALKPGASKTGGWRLSDLSASGTFVRENILGTKDSTKDIPETPAKDKPTQ 350
M2UBG4|cellulase|EC                 SALKPGASKTGGWRLSDLSASGTFVRENILGTKDSTKDIPETPSKDKPTQ 350
H6WUK0|cellulase|EC                 SALKPGASKTGGWPLSDLTASGTFVRENIRGTKDSTKDSPETPAQDNPIQ 350
                                    ************* ****:********** ******** ****::*:* *

fig|6666666.106072.peg.3365_Ba      ENGISVQYRAGDGSMNSNQIRPQLQIKNNGNTTVDLKDVTARYWYKAKNK 400
```

```
M2UBG4|cellulase|EC                ENGISVQYRAGDGSMNSNQIRPQLQIKNNGNTTVDLKDVTARYWYKAKNK 400
H6WUK0|cellulase|EC                EKGISVQYKAGDGRVNSNQIRPQLHIKNNGNATVDLKDVTARYWYNVKNK 400
                                   *:******:**** :*********:******:*************:.*** 

fig|6666666.106072.peg.3365_Ba     GQNVDCDYAQIGCGNVTHKFVTLHKPKQGADTYLELGFKNGTLAPGASTG 450
M2UBG4|cellulase|EC                GQNFDCDYAQIGCGNVTHKFVTLHKPKQGADTYLELGFKNGTLAPGASTG 450
H6WUK0|cellulase|EC                GQNFDCDYAQIGCGNLTHKFVTLHKPKQGADTYLELGFKTGTLSPGASTG 450
                                   ***.***********:**********************.***:****** 

fig|6666666.106072.peg.3365_Ba     NIQLRLHNDDWSNYAQSGDDYSFFKSNTFKTTKKITLYDQGKLIWGTEPN 500
M2UBG4|cellulase|EC                NIQLRLHNDDWSNYAQSGD-YSFFKSNTFKTTKKITLYDQGKLIWGTEPN 499
H6WUK0|cellulase|EC                NIQLRLHNDDWSNYAQSGD-YSFFQSNTFKTTKKITLYHQGKLIWGTEPN 499
                                   ****************** ****:*************.********** 
```

8.3.8 Beta-glucanase alignment to identify KEGG-KASS endo-beta-glucanase in *Bacillus subtilis* AJF2 (fig|6666666.106072.peg.1834)

```
J7JXX5|cellulase|EC                MAKLDETLTMLKDLTDAKGIPGNEREVRQVMKSYIEPFADEVTTDRLGSL 50
fig|6666666.106072.peg.1834_Ba     MAKLDETLTMLKDLTDAKGIPGNEREVRQVMKSYIEPFADEVTTDRLGSL 50
fig|6666666.106072.peg.3365_Ba     -MKRSISIFITCLLITLLTMGG-IMASPASAAGTKTPVAKNGQLSIKGTQ 48
                                     * . :: :   *     : *          .   *.*.:   .  *: 

J7JXX5|cellulase|EC                IAKKTGAENGPKIMIAGHLDEVGFMVTQ-----------ITDKGFIRFQT 89
fig|6666666.106072.peg.1834_Ba     IAKKTGAENGPKIMIAGHLDEVGFMVTQ-----------ITDKGFIRFQT 89
fig|6666666.106072.peg.3365_Ba     LVNRDGKAVQLKGISSHGLQWYGEYVNKDSLKWLRDDWGITVFRAAMYTA 98
                                   :.:: *     * : :  *:  *  *.:           **      : : 

J7JXX5|cellulase|EC                VGGWWAQVMLAQRVTIVTKKGEITGVIGSKPPHILSP-EARKKSVEIKDM 138
fig|6666666.106072.peg.1834_Ba     VGGWWAQVMLAQRVTIVTKKGEITGVIGSKPPHILSP-EARKKSVEIKDM 138
fig|6666666.106072.peg.3365_Ba     DGGYIDNPSVKNKVKEAVEAAKELGIYVIIDWHILNDGNPNQNKEKAKEF 148
                                    **:  :  : ::*. ..: .:  *:      ***.  :..::. : *:: 

J7JXX5|cellulase|EC                FIDIG-------------ASSREEALEWG---------------------- 154
fig|6666666.106072.peg.1834_Ba     FIDIG-------------ASSREEALEWG---------------------- 154
fig|6666666.106072.peg.3365_Ba     FKEMSSLYGNTPNVIYEIANEPNGDVNWKRDIKPYAEEVISVIRKNDPDN 198
                                   * ::.             *.. :  ::* 

J7JXX5|cellulase|EC                -------------------VLPGDMIVPHFEFTVMNNEKFLLAKA----- 180
fig|6666666.106072.peg.1834_Ba     -------------------VLPGDMIVPHFEFTVMNNEKFLLAKA----- 180
fig|6666666.106072.peg.3365_Ba     IIIVGTGTWSQDVNDAADDQLKDANVMYALHFYAGTHGQFLRDKANYALS 248
                                                     * .  ::  :.* . .: :**  ** 

J7JXX5|cellulase|EC                ----------------------------WDNRIGCAIAIDVLRNLQNTDH 202
fig|6666666.106072.peg.1834_Ba     ----------------------------WDNRIGCAIAIDVLRNLQNTDH 202
fig|6666666.106072.peg.3365_Ba     KGAPIFVTEWGTSDASGNGGVFLDQSREWLKYLDSKTISWVNWNLSDKQE 298
                                                              * : :..     *  **.:.:. 

J7JXX5|cellulase|EC                PNIVYGVGTVQEEVGLRGAKTAAHTIQPDIAFGVDVGIAGDTPGISEKEA 252
fig|6666666.106072.peg.1834_Ba     PNIVYGVGTVQEEVGLRGAKTAAHTIQPDIAFGVDVGIAGDTPGISEKEA 252
fig|6666666.106072.peg.3365_Ba     SSSALKPGASKTGGWRLSDLSASGTFVRENILGTKDSTKDIPETPAKDKP 348
                                   .. .   *: :      .  :*: *:  :  :*.. .  . .   ::.:. 

J7JXX5|cellulase|EC                QSKMGKGPQIIVYDASMVSHKG-------LRDAVVATAEEAGIPYQFDAI 295
fig|6666666.106072.peg.1834_Ba     QSKMGKGPQIIVYDASMVSHKG-------LRDAVVATAEEAGIPYQFDAI 295
fig|6666666.106072.peg.3365_Ba     TQENGISVQYRAGDGSMNSNQIRPQLQIKNNGNTTVDLKDVTARYWYKAK 398
                                    .: * . *  . *.** *::         .. ...  ::.   * :.* 

J7JXX5|cellulase|EC                AGG-GTDSGAIHLTANGVPALSITIATR----------------------Y 323
fig|6666666.106072.peg.1834_Ba     AGG-GTDSGAIHLTANGVPALSITIATR----------------------Y 323
fig|6666666.106072.peg.3365_Ba     NKGQNVDCDYAQIGCGNVTHKFVTLHKPKQGADTYLELGFKNGTLAPGAS 448
                                     * ..*..  :: ...*.   :*: . 

J7JXX5|cellulase|EC                IHTHAAMLHRDDYENAVKLIT--EVIKKLDRKTVDEITYQ---------- 361
fig|6666666.106072.peg.1834_Ba     IHTHAAMLHRDDYENAVKLIT--EVIKKLDRKTVDEITYQ---------- 361
fig|6666666.106072.peg.3365_Ba     TGNIQLRLHNDDWSNYAQSGDDYSFFKSNTFKTTKKITLYDQGKLIWGTE 498
                                     .    **.**:.* .:     ..:*.   **..:** 

J7JXX5|cellulase|EC                --
fig|6666666.106072.peg.1834_Ba     --
fig|6666666.106072.peg.3365_Ba     PN 500
```

8.3.9 *Bacillus subtilis* AJF2 licheninase (EC 3.2.1.73) alignment with top hit from PSI-BLAST search

Amino acid sequence alignment of fig|6666666.106072.peg.2457 (*B. subtilis* AJF2 licheninase) with most homologous gene A3DJ79.

```
Licheninase_EC3.2.1.73_Bacillu     MPYLKRVLLLLVTGLFMSLFAVTSTASAQTGGSFFDPFNGYNSGFWQKAD 50
UniRef50_A3DJ79_Glycoside_hydr     ---------------SSTNPAPTKNPNEEWRLVWSDEFNGINMANWSYDD 35
```

```
                                                :  * *.... :    : * *** * . *.  *

Licheninase_EC3.2.1.73_Bacillu      GYSNGNMFNCTWRANNVSMTSLGEMRLALTSPAYNKFDCGENRSVQTYGY 100
UniRef50 A3DJ79 Glycoside hydr      GRWNGEVQSYTQNN---AYIKDGALVIEARKEDTYHYTSSKLITKGKWKY 82
                                    *  **:: . * .    :  . * : :   .    :: ..:  :  .: *

Licheninase_EC3.2.1.73_Bacillu      GLYEVRMKPAKNTGIVSSFFTYTGPTDGTPWDEIDIEFLGKDTTKVQFNY 150
UniRef50_A3DJ79_Glycoside_hydr      GKFEIRAKMPQGQGIWPAIWMMPEDGTWPKCGEIDMELLGHEPDKIYGTI 132
                                    * :*:* * .:. ** .:::  .     .  .***:*:**::. *:  .

Licheninase_EC3.2.1.73 Bacillu      YTNGAGNHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQLKHTATNQI 200
UniRef50_A3DJ79_Glycoside_hydr      HFGEPHKESQGTYTGQTFADDFHVYSIEWEPGEIRWYIDGKLYHVANDWP 182
                                    : . . :..: .  *   *: :*.*:::*:*..*:**:**:* *.*.:

Licheninase_EC3.2.1.73_Bacillu      PTTPGKIMMNLWNGTGVDEWLGSYNGVNPLYAHYDWVRYTKK 242
UniRef50_A3DJ79_Glycoside_hydr      FDQNFFLILNISVGGGWPGYPDETT-VFPQQMVVDYVRVYQK 223
                                          :::*:  * *   : .. . * *     *:**  :*
```

8.3.10 *L. rhizovicinus* AEOb xylanase from contig 42 alignment

Alignment with three top hits from NCBI protein BLAST:

```
fig|6666666.106052.peg.2773         --MHLVRIALLAATFFIACPGNA-ATRSPTDPSAPVSIDLWPSTPPGDGP 47
gi|738154750|ref|WP_036112109.      --MSLLRVALLAAAIGLACTGQV-AAQAAPDAGKPTSIDLWPGSPPGGGS 47
gi|748174891|ref|WP_039748441.      -MASISRACRAFAILVLACGAGP-AVAAAG--AAPTLIPLWPDAVLADAP 46
gi|475007423|emb|CCP11714.1|        MSLHVGKLRGLAVLLFLLNPLLAGAVPAAPPASAPEVIPLWPQDHWQQRA 50
                                        : :     . : :       *. :.   . *  * ***       .

fig|6666666.106052.peg.2773         GPRGPEHVYASGSHSGAVSNISVPRIEIYTPAHPNGTAIIVIGGGGYARI 97
gi|738154750|ref|WP_036112109.      GPKGPERIGQSGTGVGAVSNVSRPRIEVYRPARPNGAAVILIGGGGYFRI 97
gi|748174891|ref|WP_039748441.      -VRGPERVGRDGKGTGAISNIRTPRLELYRPATPNGTAVLIAGGGGYFRI 95
gi|475007423|emb|CCP11714.1|        --RGPERIGAEGSAKGAVSGISNARMEVYRPLKPNGTAVVIFGGGGYFRI 98
                                      :***::  .*.  **:*.:  .*:*:* *  ***:*::: ***** **

fig|6666666.106052.peg.2773         GMGHEVMPTAAWLETQGVTAVVLYYRLPADHWPAVAPFQDAQRAVRVLRA 147
gi|738154750|ref|WP_036112109.      GIGHEVIPTAKWLAALGVTPVVLYYRLPADHWPAVAPFQDAQRAIRVLRA 147
gi|748174891|ref|WP_039748441.      QIGNEAVPTAKWLAALGVTAAVLYYRLPADGWPASAPFADAQRAMRILRS 145
gi|475007423|emb|CCP11714.1|        QIGSAAKPTAQWLQSLGVTAAVVYYRLPADGWAPVSPFQDGQRAMRLLRA 148
                                     :*  . *** ** : ***..*:******* *.. :** *.***:*:**:

fig|6666666.106052.peg.2773         HASQFGINPARIGVLGFSAGGNLAGIAATRSAAAFYRPVDAADQLSARPD 197
gi|738154750|ref|WP_036112109.      HASELGIDPKRVGVVGFSAGGNLAGIAETRFADAFYPAVDAADQQSARPD 197
gi|748174891|ref|WP_039748441.      RAGELGIDPRQIGVLGMSAGGHLAAITETRYADAFHEPLDAIDQVSARPD 195
gi|475007423|emb|CCP11714.1|        QAASLGVDPERIGALGFSAGGNLAGIVATRFDAPLYPPVDAADQLSARPD 198
                                    :*..:*::* ::*.:*:****:**.*. **   .:: .:** ** *****

fig|6666666.106052.peg.2773         FAALLYPVISLEQPYDTTRSGRELSTQADAAQAYSVQLHVTHNAPPTFLA 247
gi|738154750|ref|WP_036112109.      FAGLIYPVISLQKPFDTTRSSRELSTQSDAVQAYSVELHVTHDTPPTFLA 247
gi|748174891|ref|WP_039748441.      FAALIFPVISLQAPLDTTRSKRELSTQKDAVEAYSAEMHVSHDTPPTFLA 245
gi|475007423|emb|CCP11714.1|        FLGMIYPVVSMRPPLDHTRSKRELASQNDAQQAYSVEQHVRKDMPPVFLA 248
                                    * .:::**:*:. * * *** ***::* ** :***.: ** :: **.***

fig|6666666.106052.peg.2773         QAKDDPIANVGNSQIMYDALKKNGVPAELHVFDMGGHGWGMGTPGSEVAA 297
gi|738154750|ref|WP_036112109.      QAADDPIANIGNSLVMFDALHKNDVDTEMHVFDKGGHGWGLGDPGSAPAQ 297
gi|748174891|ref|WP_039748441.      HAADDPIADVGHSLAMFNALHAASVPAELHVFEAGGHGWGLGAPSTLVAQ 295
gi|475007423|emb|CCP11714.1|        QAVDDRTVDVGHSVAMYQATHAAGIPVELHLFEAGGHSWGLGKPGTLVAQ 298
                                    :* **  .::*:*  *::* :  .: .*:*:*: ***.**:* *.:  *

fig|6666666.106052.peg.2773         WPGLFAAWAQRGGFL-------------------------------- 312
gi|738154750|ref|WP_036112109.      WPRLFAAWARRAGFFNAVDTGPFAVPATSPSNAAPAASDAIPDDN--- 342
gi|748174891|ref|WP_039748441.      WPRLFATWARSHGFM-------------------------------- 310
gi|475007423|emb|CCP11714.1|        WPRLFANWARSHGYLGPPPVRLPPHDASAIPAAAAQPPRSLDDDGAGD 346
                                    ** *** **:  *::
```

Alignment with equivalent xylanase from genome *L. rhizovicinus* DSM 16549:

```
fig|1440763.3.peg.3071_DSM_16549  MHRVRIALLAASFFVACTGHAATRGAAEPTSPVSIDLWPSTPPGEGPGPRGPERVSTSGA 60
fig|6666666.106052.peg.2773_AE0b  MHLVRIALLAATFFIACPGNAATRSPTDPSAPVSIDLWPSTPPGDGPGPRGPEHVYASGS 60
fig|6666666.106055.peg.3173_AEZa  MHLVRIALLAATFFIACPGNAATRSPTDPSAPVSIDLWPSTPPGDGPGPRGPEHVYASGS 60
                                    ** ********:**:** *.****. ::*::************:*********:* :**:

fig|1440763.3.peg.3071_DSM_16549  VSNV----SVPRIEIYSPAHPNGTAIIVIGGGGYARIGMGHEVMPTAAWLETQGVTAVVL 116
fig|6666666.106052.peg.2773_AE0b  HSGAVSNISVPRIEIYTPAHPNGTAIIVIGGGGYARIGMGHEVMPTAAWLETQGVTAVVL 120
fig|6666666.106055.peg.3173_AEZa  HSGAVSNISVPRIEIYTPAHPNGTAIIVIGGGGYARIGMGHEVMPTAAWLETQGVTAVVL 120
                                     * .    ********:******************************************

fig|1440763.3.peg.3071_DSM_16549  YYRLPADHWPAVAPFQDAQRAIRVLRAHAQQLGINPARVGVLGFSAGGNLAGIAETRSAA 176
fig|6666666.106052.peg.2773_AE0b  YYRLPADHWPAVAPFQDAQRAVRVLRAHASQFGINPARIGVLGFSAGGNLAGIAATRSAA 180
fig|6666666.106055.peg.3173_AEZa  YYRLPADHWPAVAPFQDAQRAVRVLRAHASQFGINPARIGVLGFSAGGNLAGIAATRSAA 180
                                  ********************:*******.*:******:*************** *****

fig|1440763.3.peg.3071_DSM_16549  SFYRPVDAADQLSARPDFAALLYPVISLEQPYDTTRSGRELSTQADAAQAYSVQLHVTRN 236
fig|6666666.106052.peg.2773_AE0b  AFYRPVDAADQLSARPDFAALLYPVISLEQPYDTTRSGRELSTQADAAQAYSVQLHVTHN 240
fig|6666666.106055.peg.3173_AEZa  AFYRPVDAADQLSARPDFAALLYPVISLEQPYDTTRSGRELSTQADAAQAYSVQLHVTHN 240
                                  :*********************************************************:*
```

```
fig|1440763.3.peg.3071_DSM_16549   APPTFLAQAKDDPIANIGNSQIMYEALKKNGVPAELHVFDTGGHGWGMGTPGSEVAAWPG 296
fig|6666666.106052.peg.2773_AE0b   APPTFLAQAKDDPIANVGNSQIMYDALKKNGVPAELHVFDMGGHGWGMGTPGSEVAAWPG 300
fig|6666666.106055.peg.3173_AEZa   APPTFLAQAKDDPIANVGNSQIMYDALKKNGVPAELHVFDMGGHGWGMGTPGSEVAAWPG 300
                                   ****************:*******:*************** ******************

fig|1440763.3.peg.3071_DSM_16549   LFATWARRNGFL 308
fig|6666666.106052.peg.2773_AE0b   LFAAWAQRGGFL 312
fig|6666666.106055.peg.3173_AEZa   LFAAWAQRGGFL 312
                                   ***:**:* ***
```

8.3.11 *L. rhizovicinus* AEOb xylanase from contig 489 alignment

Alignment with four homologous protein sequences from NCBI BLASTp search:

```
gi|657884926|ref|WP_029589220.      MGASMGKIMGLAAAMLLAIGGASARAAEPTTVSPEAPWRPGEGLSQIPLW 50
gi|522213754|ref|WP_020720866.      -----------MRSLIFALCVVSASS---SLFAQKRGWQPSPGHTQTPIW 36
fig|6666666.106052.peg.3024         --------------------------------------------------
gi|343779704|gb|AEM52257.1|         --------MLIVRARLLMGSLLLAMVPFATTKAAPSTWHPPQGTTELPLW 42
gi|797103989|ref|WP_045828656.      --------------------------MDGLQRRP--WQPGAPGIEVPLW 21


gi|657884926|ref|WP_029589220.      P---PNLAIARPALTGPETARVGGGIAG-RPVTMIENVERPTMTVYPAKG 96
gi|522213754|ref|WP_020720866.      PGIPPDAQFGPPANT--ETTKPGE----------VDNVSRPTMTVYSPKG 74
fig|6666666.106052.peg.3024         -------------MTGAESVRDDA---------IVENVTRPTMTIFPPKG 28
gi|343779704|gb|AEM52257.1|         P---QGTVSAPPKLKGPEQVSEVTSEASGEHWMMLQNVAVPTLTVFPPKG 89
gi|797103989|ref|WP_045828656.      P---TGMAITPPAYPGPEVFGFSKGTIAGKPIQMVQHVSRPTMAIYPAKG 68
                                                *               :::*  **::::..**

gi|657884926|ref|WP_029589220.      ANTGAAVLVFPGGGYRALAIDLEGSEVCDWLTPKGVTCVVLKYRVPGSGH 146
gi|522213754|ref|WP_020720866.      KNTGAAVVVFPGGGYFALAIDLEGTEVCDWLTSKGITCVLLKYRVPNTGP 124
fig|6666666.106052.peg.3024         PGTGTTVVVFPGGGFRVLAMDGEGTEICDWVTSKGITCVLLKYRVPTSGP 78
gi|343779704|gb|AEM52257.1|         PATGTAVMVVPGGGYRVLAMDLEGSEICDWLAAQGITCALLKYRVPASGP 139
gi|797103989|ref|WP_045828656.      VNSGAALLVFPGGGYRVLAIDLEGTEICDWATAKGITCAVLKYRVPGSGP 118
                                      :*:::::*.****: .**:* **:*:*** :.:*:**.:****** :*

gi|657884926|ref|WP_029589220.      YWNVACKCHKIPQVPMALQDAQRAMVILRHRATEWGVDPRKIGVLGFSAG 196
gi|522213754|ref|WP_020720866.      HWDGQCKCQMNPKAPMALEDAQRTVGLARFHAAEYHIDPHKIGVLGFSAG 174
fig|6666666.106052.peg.3024         QWDGSCDCRREPTVHMALQDAQRAIRLIRQQASSLHVDPHKVGVIGFSAG 128
gi|343779704|gb|AEM52257.1|         NWDPECNCRDIPKVPMALQDAQRAMGLLRAQAGRWSINPKRIGVIGFSAG 189
gi|797103989|ref|WP_045828656.      YWNDECHCRRDPPAPMALQDAQRAMGLLRERAAQFRFNPARVGVVGFSAG 168
                                     *:  *.*:  * . ***:****:: : * :*    .:* ::**:*****

gi|657884926|ref|WP_029589220.      GHLVAAV-TNADSLSYAPVDASDLESSRPDFGVALYPGHLWSG--TATDL 243
gi|522213754|ref|WP_020720866.      GHLVTDISTHFDKRSYASVDAADRESCRPNFAVALYPGHLWLSSYKKFEL 224
fig|6666666.106052.peg.3024         GNMVADV-SNASGKNYAPVDEADTLSARPDFAMALYPGRLWEG--KGVVL 175
gi|343779704|gb|AEM52257.1|         GHVAAGL-STHALRSYAPVDAADEQPSRPDFAMVMYSGHLWAGHGKGLSL 238
gi|797103989|ref|WP_045828656.      GRMVADI-SNHPARAYTPVDAADRLPSRPDFAMALYPGHLWDGEPGGIAL 217
                                    *.:.: : :      *:.** :*  ..**:*.:.:*.*:** .      *

gi|657884926|ref|WP_029589220.      YPFNKISANTPPTFIVQAQDDPVDDVRNAITYFLALRSAGAPVEMHIYAH 293
gi|522213754|ref|WP_020720866.      NPDVAVTRQTPPTFLLQAENDPVDSVNNSLVYYIALKNAGVPVEMHLYAE 274
fig|6666666.106052.peg.3024         NPTLHITAKTPPTFLVQAEDDPVDNVRNSVTYFLALKEHGVPVEMHLYAT 225
gi|343779704|gb|AEM52257.1|         VRDITVDDKVPPTFIAQATDDRTDDVRESLAYYRALVDAGVPVEMHLFAR 288
gi|797103989|ref|WP_045828656.      TKEVAIDPKAPPTFIAQAGDDATDDVRESLTYYLALQQASVPVEMHLYPE 267
                                         :  :.****: ** :* .*.*.:::.*: ** . ..*****::.

gi|657884926|ref|WP_029589220.      GGHAFALRPTAQPITRWPALVEIWLHTIGVL------- 324
gi|522213754|ref|WP_020720866.      GKHAFGLRPTKFPITRWPQLVETWLRTIGIISK----- 307
fig|6666666.106052.peg.3024         GGHAFALRKKDMPIGAWPQLAEKWMQSIGMLPAG---- 259
gi|343779704|gb|AEM52257.1|         GGHAFGLRLKDAPVAAWPQLAERWMQDIGMLPRSL--- 323
gi|797103989|ref|WP_045828656.      GGHAFGVRPTDKPISHWTGLAEGWMLRLGILSRGVPTR 305
                                    * ***.:* .  *:  *. *.* *:  :*::
```

Alignment with equivalent xylanase gene in *L. rhizovicinus* DSM 16549:

```
fig|6666666.106052.peg.3024_AE0b   ------------------------------------------MTGAESVRDDAIVENVTR 18
fig|6666666.106055.peg.417_AEZa    ------------------------------------------MTGAESVRDDAIVENVTR 18
fig|1440763.3.peg.1480_DSM_16549   MIARPPVTGPESVKNGATIENVTRPTMTVVPARGPSTGTTVVVFPGGGFRVLAMDGEGTE 60
fig|6666666.106056.peg.1826_AFDa   MAASASPQDSPASATAPATWQPPAPLTQVPLWPEGHAIARPPVTGAESVRDDAIVENVTR 60
                                                                    :  . ..*  *:  : *.

fig|6666666.106052.peg.3024_AE0b   PTMTIFPPKGPGTGTTVVVFPGGGFRVLAMDGEGTEICDWVTSKGITCVLLKYRVPTSGP 78
fig|6666666.106055.peg.417_AEZa    PTMTIFPPKGPGTGTTVVVFPGGGFRVLAMDGEGTEICDWVTSKGITCVLLKYRVPTSGP 78
fig|1440763.3.peg.1480_DSM_16549   ICDWLTSKGITCVLLKYRVPNSGPQWDGSCDCRREPTVHMALQDAQRAIRLIRQQASSLH 120
fig|6666666.106056.peg.1826_AFDa   PTMTIFPPKGPGTGTTVVVFPGGGFRVLAMDGEGTEICDWVTSKGITCVLLKYRVPTSGP 120
                                      :       .  .  *  .*     : * .     . . ...  .: *  :  :*

fig|6666666.106052.peg.3024_AE0b   QWDGSCDCRREPTVHMALQDAQRAIRLIRQQASSLHVDPHKVGVIGFSAGGNMVADVSNA 138
fig|6666666.106055.peg.417_AEZa    QWDGSCDCRREPTVHMALQDAQRAIRLIRQQASSLHVDPHKVGVIGFSAGGNMVADVSNA 138
fig|1440763.3.peg.1480_DSM_16549   VDPHKVGVIGFSAGGN-----------------------------------MVADVSNA 144
fig|6666666.106056.peg.1826_AFDa   QWDGSCDCRREPTVHMALQDAQRAIRLIRQQASSLHVDPHKVGVIGFSAGGNMVADVSNA 180
                                      .       :                                        ********

fig|6666666.106052.peg.3024_AE0b   SGKNYAPVDEADTLSARPDFAMALYPGRLWEGKGVVLNPTLHITAKTPPTFLVQAEDDPV 198
```

```
fig|6666666.106055.peg.417_AEZa    SGKNYAPVDEADTLSARPDFAMALYPGRLWEGKGVVLNPTLHITAKTPPTFLVQAEDDPV 198
fig|1440763.3.peg.1480_DSM_16549   SGKNYAPVDEADTLSARPDFAMALYPGRLWEGKGVVLNPTLHITAKTPPTFLVQAEDDPV 204
fig|6666666.106056.peg.1826_AFDa   SGKNYAPVDEADTLSARPDFAMALYPGRLWEGKGVVLNPTLHITAKTPPTFLVQAEDDPV 240
                                   ************************************************************

fig|6666666.106052.peg.3024_AE0b   DNVRNSVTYFLALKEHGVPVEMHLYATGGHAFALRKKDMPIGAWPQLAEKWMQSIGMLPA 258
fig|6666666.106055.peg.417_AEZa    DNVRNSVTYFLALKEHGVPVEMHLYATGGHAFALRKKDMPIGAWPQLAEKWMQSIGMLPA 258
fig|1440763.3.peg.1480_DSM_16549   DNVRNSVTYFLALKEHGVPVEMHLYATGGHAFALRKKDMPIGAWPQLAEKWMQSIGMLPA 264
fig|6666666.106056.peg.1826_AFDa   DNVRNSVTYFLALKEHGVPVEMHLYATGGHAFALRKKDMPIGAWPQLAEKWMQSIGMLPA 300
                                   ************************************************************

fig|6666666.106052.peg.3024_AE0b   G        259
fig|6666666.106055.peg.417_AEZa    G        259
fig|1440763.3.peg.1480_DSM_16549   G        265
fig|6666666.106056.peg.1826_AFDa   G        301
                                   *
```

8.3.12 *L. rhizovicinus* AEOb Xylanase (fig|6666666.106052.peg.3024) alignment with top hit from Phyre2 Intensive search

```
fig_6666666_106052_peg_3024_Xy     MTGAESVRDDAIVENVTRPTMTIFPPKGPGTGTTVVVFPGGGFRVLAMDG 50
UniRef50_B3VP83_Lipase/esteras     LAGEPEIENKKIAGPGGEIPVRIYTPKGEGPFPALVYYHGGGWVIGDLET 50
                                   ::*  .:.:. *.    . .: *:.*** *. .::* : ***: :  ::

fig_6666666_106052_peg_3024_Xy     EGTEICDWVTSKGITCVLLKYRVPTSGPQWDGSCDCRREPTVHMALQDAQ 100
UniRef50_B3VP83_Lipase/esteras     VDVPCRLLTNLANCVVVSVDYRLAP------------EHKFPAAAEDSY 87
                                    ..     ..  . . * :.**:..             * ..  * :*:

fig_6666666_106052_peg_3024_Xy     RAIRLIRQQASSLHVDPHKVGVIGFSAGGNMVADVSNASGKNYAPVDEAD 150
UniRef50 B3VP83 Lipase/esteras     AAAKWVAENAASIGVDPNRIAVGGDSAGGNLAAVVALMARDKREISIAYQ 137
                                    * : : ::*:*: ***:::.* * *****:.* *:  : .:       :

fig_6666666_106052_peg_3024_Xy     TLSARPDFAMALYPGRLWEGKGVVLNPTLHITAKTPPTFLVQAEDDPVDN 200
UniRef50_B3VP83_Lipase/esteras     MLIYPVTWFWNHYLRNEEDGKNPYASPLQADLSGLPPALVLTGEFDPLRD 187
                                    *     :    *  .  :**.   .*     :  **:::: .* **: :

fig 6666666 106052 peg 3024 Xy     VRNSVTYFLALKEHGVPVEMHLYATGGHAFALRKKDMPIGAWPQLAEKWM 250
UniRef50_B3VP83_Lipase/esteras     EGE--AYAERLKEAGVPVEAKRYDGMIHGFWMP------GVLEQGRKSIE 229
                                     :  :*   *** ***** : *    *.* :       *.  *  :.

fig_6666666_106052_peg_3024_Xy     QSIGMLPAG- 259
UniRef50 B3VP83 Lipase/esteras     QAADALKRVF 239
                                   *: . *
```

8.3.13 *Bacillus licheniformis* AHY1 beta-xylosidase alignment with top hits

8.3.13.1 Beta-xylosidase fig|6666666.106061.peg.3748 and Q65MB6 alignment

```
fig6666666106061peg3748_Beta-x     MSGEHTYTNPVLTGFHPDPSIIRVGEDYYMVNSTFQYFPAIVISHSKDLV 50
UniRef50_Q65MB6_Glycoside_Hydr     MSGEHTYTNPVLTGFHPDPSIIRVGEDYYMVNSTFQYFPAIVISHSKDLV 50
                                   **************************************************

fig6666666106061peg3748 Beta-x     HWKIIGHGITENEGLDLSDINDSHGIWAPDISYHNGTFYIFATHRLNGPT 100
UniRef50_Q65MB6_Glycoside_Hydr     HWKIIGHGITENEGLDLSDINDSHGIWAPDISYHNGTFYIFATHRLNGPT 100
                                   **************************************************

fig6666666106061peg3748_Beta-x     VVNGRKLIRRQIMIKSSRPEGPYSKPVFIDEGSGIDPSHFVDGDGKHYML 150
UniRef50_Q65MB6_Glycoside_Hydr     VINGRKLIRRQIMIKSSRPEGPYSKPVFIDEGSGIDPSHFVDGDGKHYML 150
                                   *:**************************************************

fig6666666106061peg3748_Beta-x     LSPACTLFPLNEECTDISGEPVQIWEGTGRRAPEGPHLLKKDRYYYAILA 200
UniRef50_Q65MB6_Glycoside_Hydr     LSPACTLFPLNEECTDISGEPVQIWEGTGRRAPEGPHLLKKDGYYYAILA 200
                                   ******************************************* *******

fig6666666106061peg3748 Beta-x     EGGTGYSHSITTARSTHLYGPYEPCPYNPILTQTDPDAPIQRAGHGKLVE 250
UniRef50 Q65MB6 Glycoside Hydr     EGGTGYSHSITTARSTHLYGPYEPCPYNPILTQTDPDAPIQRAGHGKLVE 250
                                   **************************************************

fig6666666106061peg3748_Beta-x     TQNGEWWAVYLCGRPNQGSYTTVGRETALDPVEWTDDGWFVINSLKGPSL 300
UniRef50_Q65MB6_Glycoside_Hydr     TQNGEWWAVYLCGRPNQGSYTTVGRETALDPVEWTDDGWFVINNLKGPSL 300
                                   *******************************************.******

fig6666666106061peg3748_Beta-x     VQRAPNLPQVKWDEKNFDDFDEDTLGLDWQFVRNPDHSGWSLTERPGYLR 350
UniRef50_Q65MB6_Glycoside_Hydr     VQRAPNLPQVKWDEKNFDDFDEDTLGLDWQFVRNPDHSSWSLIERPGYLR 350
                                   ************************************.*** *******

fig6666666106061peg3748_Beta-x     LWTGDWDLHDIRAKNTVVRREKHHLYSAGVKLDFSPSASGEQAGIVCYYS 400
UniRef50 Q65MB6 Glycoside Hydr     LWTGDWDLHDIRAKNTVVRREKHHLYSAGVKLDFSPSASGEQAGIVCYYS 400
                                   **************************************************

fig6666666106061peg3748_Beta-x     TNNYLKCCLIYEEGLKIKVVENRSGCQKTLGKKHAEAGPLFLKAVINKQK 450
UniRef50_Q65MB6_Glycoside_Hydr     TNNYLKCCLIYEEGLKIKVVENRSGCQKTLGKKHAEAGPLFLKAVINKQK 450
                                   **************************************************
```

```
fig6666666106061peg3748_Beta-x      RDFYYSYEGKHWHHAGGTEDASFLSDEGSRDAKGHTGTMVGIFANNGGSG 500
UniRef50_Q65MB6_Glycoside_Hydr      RDFYYSYEGKHWHHAGGTEDASFLSDEGSRDAKGHTGTMVGIFANNGGSG 500
                                    **************************************************

fig6666666106061peg3748_Beta-x      RKAAADFDWFRYIAY 515
UniRef50_Q65MB6_Glycoside_Hydr      RKAAADFDWFRYIAY 515
                                    ***************
```

8.3.13.2 Beta-xylosidase fig|6666666.106061.peg.3747 and Q65MB7 alignment

```
fig6666666106061peg3747 Beta-x      MAKSLTGKAKHHPPVWNPNIDGGRFCNPVLFMDYSDPDVIRKGSDFFMAA 50
UniRef50_Q65MB7_Glycoside_Hydr      MAKSLTGKAKHHPPVWNPNIDGGRFCNPVLFMDYSDPDVIRKGSDFFMAA 50
                                    **************************************************

fig6666666106061peg3747 Beta-x      SSFSCFPGIPILHSTDLVNWSLISHVFDELPFGEYEAPAHGKGAWAPSIR 100
UniRef50 Q65MB7 Glycoside Hydr      SSFSCFPGIPILHSTDLVNWSLISHVFDELPFGEYEAPAHGKGAWAPSIR 100
                                    **************************************************

fig6666666106061peg3747_Beta-x      YHNGEFWVYFATPDEGIFMSKTSDPFKGWEPLIQVKKTKGWIDPCPFWDD 150
UniRef50_Q65MB7_Glycoside_Hydr      YHNGEFWVYFATPDEGIFMSKTSDPFKGWEPLIQVKKTKGWIDPCPFWDD 150
                                    **************************************************

fig6666666106061peg3747 Beta-x      DGNAYLIHAYANSRIGIKSKLNLCRMKKDGTALEDDGEIVFDGTLNHPTI 200
UniRef50 Q65MB7 Glycoside Hydr      DGNAYLIHAYANSRIGIKSKLNLCRMKKDGTALEDDGEIVFDGTLNHPTI 200
                                    **************************************************

fig6666666106061peg3747_Beta-x      EGPKLYKKDGYYYIFAPAGGVKTGWQTVLRSKQIYGPYEDKVVLHQGNTV 250
UniRef50 Q65MB7 Glycoside Hydr      EGPKLYKKDGYYYIFAPAGGVKTGWQTVLRSKQIYGPYEDKVVLHQGNTV 250
                                    **************************************************

fig6666666106061peg3747_Beta-x      INGPHQGGWVELDSGESWFIHFQDRGAFGRVVQLQPVRWFDGWPLMGEDI 300
UniRef50_Q65MB7_Glycoside_Hydr      INGPHQGGWVELDSGESWFIHFQDRGAFGRVVQLQPVRWFDGWPLMGEDI 300
                                    **************************************************

fig6666666106061peg3747 Beta-x      NGDGIGEPVLHWEKPKTGAKEVKLPFDHGNDDFSSSDRLALHWQWEANPK 350
UniRef50 Q65MB7 Glycoside Hydr      NGDGIGEPVLHWEKPKTGAKEVKLPFDHGNDDFSSSDRLALHWQWEANPK 350
                                    **************************************************

fig6666666106061peg3747_Beta-x      QEWYSLSASPGRLRLYSRGIAHPFKLSRLPNLLMQKFSGPSFTATVKLSL 400
UniRef50 Q65MB7 Glycoside Hydr      QEWYSLSASPGRLRLYSRGIAHPFKLSRLPNLLMQKFSGPSFTATVKLSL 400
                                    **************************************************

fig6666666106061peg3747_Beta-x      NALSETMMAGLAVMGREYAYLGFSKSENGYRLSVYEGRIEGDQEDVKTEE 450
UniRef50_Q65MB7_Glycoside_Hydr      NALSETMMAGLAVMGREYAYLGFSKSENGYRLSVYEGRIEGDQEDVKTEE 450
                                    **************************************************

fig6666666106061peg3747_Beta-x      RGVHAEMVYLRVTVDEQAVCRFSCSMDRRHFTQIGNSFQAVPGLWTGAKA 500
UniRef50 Q65MB7 Glycoside Hydr      RGVHAEMVYLRVTVDEQAVCRFSCSMDRRHFTQIGNSFQAVPGLWTGAKA 500
                                    **************************************************

fig6666666106061peg3747_Beta-x      GVYCADTGGEDGQDGWCDIEWFSVEPAGGECHER 534
UniRef50_Q65MB7_Glycoside_Hydr      GVYCADTGGEDGQDGWCDIEWFSVEPAGGECHER 534
                                    **********************************
```

8.3.14 Cytoplasmic alpha-amylases from *B. coagulans* ABQ aligned top PSI-Blast hits

8.3.14.1 fig|34521918.peg.2080 alignment with B9E9W0

```
fig34521918peg2080_Cytoplasmic      MERNHTIMQFFEWNTPADGSHWNRLKEMAPELKKSGIDAVWLPPVTKGQS 50
UniRef50_B9E9W0_Amylase_homolo      MTKNYTMMQYFEWHANGDGRHWKRLKEDAPKLKESGIDAIWLPPACKADH 50
                                    * :*:*:**:***:: .** **:**** **:**:*****:****. *.:

fig34521918peg2080_Cytoplasmic      DMDNGYGVYDHYDLGEFDQKGTVRTKYGTKQQLHEAINACHEHDIQVYID 100
UniRef50 B9E9W0 Amylase homolo      VTNTGYSIYDLYDLGEFDQKGQVRTKYGTKEELLDAIKACHENDIKVYAD 100
                                      :.**.:** ********** ********::* :**:****:**:** *

fig34521918peg2080_Cytoplasmic      VVMNHKAGADETESFQVVEVDPMDRNKEISEPFEIEGWTKFNFTNRKGKY 150
UniRef50_B9E9W0_Amylase_homolo      IVLNHKAGADETETIKVLEVDTNDRNHVISEPFEIEAYTKFEFPGRNNKY 150
                                    :*:**********:::*:***. ***: ********.:***:*..*:.**

fig34521918peg2080_Cytoplasmic      SDFTWNHTHFSGVDYDNRTGRNGIFRIVGENKHWSEHVDNEFGNFDYLMY 200
UniRef50_B9E9W0_Amylase_homolo      SDFKWNHTHFNGTDYDHKTGRKGIFKIIGENKDWNEFVDDENGNFDYLMF 200
                                    ***.******.*.***::***:***:*:****.*.*.**:* *******:

fig34521918peg2080_Cytoplasmic      ADIDYNHPDVKKEMIEWGKWLADTTGCDGYRLDAIKHINHDFIRDFAAAL 250
UniRef50 B9E9W0 Amylase homolo      TNIDYKHPDVRQHTIEWGKWLIDTLGIDGMRMDAVKHIESYFIKDFSDAM 250
                                    ::***:****::. ******* ** * ** *:**:***:  **:**: *:

fig34521918peg2080_Cytoplasmic      MEHRGDHFYFVGEFWNPQLEACQKYLDHVQFKIDLFDVALHYKLHEASKK 300
UniRef50_B9E9W0_Amylase_homolo      REHAGEDFYFLGEYWNADLGKNEKFLEEADYNTDLFDVKLHFNFKTASEE 300
                                     ** *:.***:**:**.:*   :*:*:..::: ***** **:::: **::

fig34521918peg2080_Cytoplasmic      GRAFDLTTIFHDTLVQTHPLNAVTFVDNHDSQPNESLESWVDDWFKQSAY 350
UniRef50 B9E9W0 Amylase homolo      GSAYDLRTLFEDTIVEKHPELAVTFVDNHDSQPGEALESFVKDWFKQSAY 350
                                    * *:** *:*.**:*:.**  ************.*:***:*.********
```

```
fig34521918peg2080_Cytoplasmic      ALILLRKDGYPCVFYGDMYGIGGDNPIPGKKDALSPLLSVRREKAYGEQD 400
UniRef50_B9E9W0_Amylase_homolo      ALILLRKDGYPCIFYGDYYGIGGEEPVEGKQLAIDPLLYIRQNKAYGEQD 400
                                    ************:**** *****::*: **: *:.*** :*::******* 

fig34521918peg2080_Cytoplasmic      DYFDHPNTIGWVRRGVPEIPHSGCAVVISNGEN-GEKRMLVGKERAGEVW 449
UniRef50_B9E9W0_Amylase_homolo      DYFDHPNVIGFVRRGNGN--KTGCAVIINSSEEEAEKQMFVGKERAGEVW 448
                                    *******.**:****  :  ::****:*...*: .**:*:********** 

fig34521918peg2080 Cytoplasmic      VDATGNRQEKVTIGEDGYAGFPVNGGSVSVWVQETDEN 487
UniRef50 B9E9W0 Amylase homolo      YDYTNTREDRITIDEEGNGLFKVNPGSVSVYCEEE--- 483
                                     * *..*::::**.*:* . * ** *****: :*
```

8.3.14.2 fig34521918peg857 alignment with B1YMN6

```
fig34521918peg857 Cytoplasmic       MEKKFFSRLSILMLSLLLVAGSISYFPKSAKAYTSGTSLDNRVIFQSFSL 50
UniRef50_B1YMN6_Dextransucrase      --MKNTKKVSAGLLATLVATSSFGVAPKQAAAYTSGEKLDNHVIFQSFSL 48
                                       *  .::*  :*: *:.:.*:.  **.* ***** .***:******** 

fig34521918peg857_Cytoplasmic_      YMPYESNMYKILSAKGSELKDWGITDIWLPPAYRSFNMARYMEGYAIADR 100
UniRef50_B1YMN6_Dextransucrase      YQPYDSNMYRTLAKKGDLLNSWGVTDVWMPPAYRSFDMARYMEGYAIADR 98
                                    * **:****: *: **. *:.**:**:*:*******:************** 

fig34521918peg857 Cytoplasmic       YDLGEFNQGPNNTRPTKYGTSDELKSMVSALHASGLKVQEDLVPNQVLGL 150
UniRef50_B1YMN6_Dextransucrase      YDLGEFPQGPGGTTATKYGKASHLEMMVDMLHDDNIKVQMDLVPNQMLGL 148
                                    ****** ***..* .****.:..*: **. ** ..:*** ******:*** 

fig34521918peg857 Cytoplasmic       GKREAVYVTRVDQNGNLFKNPYTTGLTTQIRADLYLAYTKGGGEGQAKYG 200
UniRef50 B1YMN6 Dextransucrase      NKREAVFVRRATSSGELFTNPYTGGQTTKTLATPYLAYTKGGGQGQEKYG 198
                                    .*****:* *. ..*:**.**** * **:  *  *********:** *** 

fig34521918peg857_Cytoplasmic_      YIKEWNKKYFNGTSVQGQGMDRVMKDSEGIPYRYFGPNNPKNHLPSWLNE 250
UniRef50_B1YMN6_Dextransucrase      YLKEWNKTFLNGTSLQGQGTGRVMTDKDGKPYRYFGPENEKNYLPEWLIE 248
                                    *:*****.::****:**** .***.*.:* *******:* **:**.** * 

fig34521918peg857 Cytoplasmic       AAAANKINTVDTYFSVDGWYAAKDASTSDNYWKPMLMNY--DPGYLKYMK 298
UniRef50_B1YMN6_Dextransucrase      ASKTQNLNVVDTYLAADGWYEVSPQN-----WKPMLSQYAKDPGYLAYMK 293
                                    *: ::::*.****::.**** ..  .     ***** :*  ***** *** 

fig34521918peg857 Cytoplasmic       SHGYSTVDDILN-GDNGEIASLTDAYIASQPGYGFGSEERSFKNDNSGSD 347
UniRef50 B1YMN6 Dextransucrase      ENGFETKEALLASADNGTIAKLTEEYMKTQATYGYGTEERSYQNDNSGID 343
                                    .:*:.* : :*  .*** **.**: *: :*. **:*:****::***** * 

fig34521918peg857_Cytoplasmic_      DQDQFLLFVKKNGTTLHNLNNTISGKKQFLLGMDILNGNPTVQKEQIHWMN 397
UniRef50_B1YMN6_Dextransucrase      IEDQFLFVDETGFPLQAYNKTMTNNDEFLLGVDLANSNTEVIKEQKNWMK 393
                                     :******.:.* .*:  *:*::.:.:****:*: *.*. * *** :**: 

fig34521918peg857 Cytoplasmic       WLLDTYQFDGFRIDAASHYDQQVLLDEADVMKQHFGNNLNDHLSYIETYE 447
UniRef50_B1YMN6_Dextransucrase      WMLETYKFDGFRIDAASHYDTAILKAEAQVAKEHFGK--KEHLSYIESYK 441
                                    *:*:**:************  :*  **:* *:***:  ::******:*: 

fig34521918peg857_Cytoplasmic_      SAGTNFENANGNPQLMMDYALFYSLQNALGKNSPSNNLSTIATNAVVNRA 497
UniRef50 B1YMN6 Dextransucrase      SEQKAYMKANNDEQLIMDSPLYFTMRTALGNEASKRPLSAIATGSTINRA 491
                                    *  . : :**.: **:** .*::::::.***:::... **:***.:.:*** 

fig34521918peg857_Cytoplasmic_      GAGTANATPNWSFVNNHDQEKNRVNSIMLDQYGIKPGTHYGTSTPKAFQD 547
UniRef50_B1YMN6_Dextransucrase      GNGSTDVSANWSFVNNHDQEKNRVNQIMLDLYGIKTGIQYKAGEEPSFEK 541
                                    * *:::.:.****************.**** ****.* :* :.   :*:. 

fig34521918peg857_Cytoplasmic_      LYDKKTEAKALDIYEKDMESTVKKYAPSNVPSQYAYVLTNKDTVPTVFYG 597
UniRef50 B1YMN6 Dextransucrase      MYDKDTEKQALGIYNAELASTKKKYSVDNVVSQYAFLLTNKDTVPTVYYG 591
                                    :***.** :**.**: :: ** ***: .** ****::**********:** 

fig34521918peg857_Cytoplasmic_      DLYKTNASYMSERTPYYDTIVKLLKVRKNYAYGNQQVTNYKSNTSGTAGK 647
UniRef50_B1YMN6_Dextransucrase      DLYQTDASYMSKQTPYYDEITNLLKVRKQYAYGAQKVAYHTTNTSKVAGS 641
                                    ***:*:*****::***** *.:******:**** *:*: :.:*** .**. 

fig34521918peg857_Cytoplasmic_      DLISSVRYGNDRNTGVATVIGNNPKTDTTIKVNMGSRHANQTFEDATGFH 697
UniRef50_B1YMN6_Dextransucrase      HLISSVRLGKDRNTGVATVIGKKSTLNTTIKVDMGKQHTNQVFVDASGVT 691
                                    .****** *:***********::.. :*****:**.:*:**.* **:*. 

fig34521918peg857_Cytoplasmic_      NEKLVTDSKGILTVHVKGTQNARVKGYLGVWIPAKKAATPKQGPALKYGK 747
UniRef50 B1YMN6 Dextransucrase      QTKLVTDKNGILTVPVKGMRTAEVNGYLGVFVPQTTKAP----------- 730
                                    : *****.:***** *** :.*.*:*****::* .. *. 

fig34521918peg857_Cytoplasmic_      YVTVTNKHYAVYQDFNWKKKNVNAVNKTYLAKVQYHHSNGSTYLSLYDGK 797
UniRef50_B1YMN6_Dextransucrase      --TATIEKASVYQGKGINLKTKVLNTSSTVSSIRYQVAD--TSKATVDTT 776
                                      *.* :: :***. . : *.    ..: ::.::*: ::  *  :  * . 

fig34521918peg857 Cytoplasmic       GKWAGYINAKAAKTGSGKQGAAIQYGKSVKVTSKNYGVYQNFNWKKKNIR 847
UniRef50_B1YMN6_Dextransucrase      GRLVGKASGKTTVTATVTLKDGFVLTTTLPIETKVNEVKLKATQKTLKKG 826
                                    *: .*  ..*:: *.: .   .:   .:: : :*   *  : . *. : 

fig34521918peg857_Cytoplasmic_      AVNKTYLAKYIYYHINGLSYLSLYDNKGKWIGYINAKAVKSK 889
UniRef50_B1YMN6_Dextransucrase      QTTTIGYAS-ATDKIKSVSYASLN----KKIATVNTKG---- 859
                                     ...   *.    :*:.:** **     * *. :*:*. 
```